

Questions for Clinton's inauguration

Next Wednesday's ceremony in Washington is bound to be an uplifting and even moving ceremony, but not all of the new president's hankerings for change are to be equally keenly awaited.

INAUGURATION days in the United States are mostly memorable occasions. They should be, for the changing of the guard in the United States must eventually affect us all. There will be an inauguration day next week, on Wednesday, when Mr Bill Clinton takes over from President George Bush. By then, Clinton will have no doubt had time, in the breathing space after November's hectic election, to string together in a brief speech several phrases that will stick in the world's mind — "there is nothing to fear but fear itself" and the like. Unlike his predecessor, Clinton has a way with words (and we must hope that his laryngitis allows him to make them heard), so that will be the easy part. What comes afterwards will be more difficult.

Even so, it matters that Clinton should strike the right note next Wednesday. The US President differs in a subtle way from the heads of other elected governments in that he (or, yet, she) is also elected leader. The president commands not merely the armed forces, but also a large measure of loyalty and support among the American people — Clinton's way of referring to his electors during the election campaign. So a new incumbent can strike out in a new direction and be reasonably sure that he will be followed, at least for a time. New presidents know that, but too often respond by promising at their inauguration that they will, in return, secure 'leadership' for the United States. Bush offered that in plenty four years ago, without quite saying what he meant.

Clinton must avoid that trap. To ask that is not to diminish either the achievement or the promise of the United States. How could that be when the surviving superpower dutifully sought United Nations sanction for its organization of the war in the Gulf in 1991, and now seems ready to be the world's policeman for a time? Or when Jefferson's doctrine of "life, liberty and the pursuit of happiness" seems, for the time being, to have infected even its previous enemies? The United States is now also the chief source of new knowledge, both in science and in other fields of scholarship; will not great things flow from that? But the world has, thank goodness, become a more complicated and thus more interesting place. Leadership is not what it needs, but more generalized cleverness. Beating Japan (and, next decade, South China) at consumer electronics might seem like leadership, but would be a distraction.

The cautious part of Clinton has not yet declared himself on science, and on scholarship generally, except to confirm his predecessors' ambition that the workforce should be well educated. True, he has made an excellent nomination for his

chief science adviser (see *Nature* 361, 2; 1993); John Gibbons, if he is given the authority, could restore the tradition that his office takes a broad view of its terms of reference — the Strategic Defense Initiative as well as the Superconducting Super Collider and the success rate of research grant applications to the National Institutes of Health. It will help that Gibbons now knows the Capitol and its people as well as he does. Strike one to Clinton.

The worry is what the new president has been saying about international trade and domestic employment under the rubric of 'competitiveness' — Washington's word for the state of US industry since President Carter's time. There is no reason why Clinton should follow Reagan and Bush (and their quondam cheerleaders in Britain) by letting the market decide how many people have useful and gainful jobs, but there are serious risks in trying to accelerate the conversion of seed-corn into full-grown corn plants, not the least of which is the temptation to cut back on the seed producers, or to keep telling them what kinds of seeds to produce. One of the achievements of the past decade in the United States has been a substantial growth of the basic research enterprise. With patience, this will yield profit and prosperity. Impatience, and an indiscriminating zest for change, could hazard that. □

Rothschild revisited?

Industry's record on research casts doubt on pleas to fund British science on a 'purchaser-provider' basis.

WHEN in 1971 the late Lord Rothschild proposed putting publicly funded applied research on a 'customer-contractor' basis, he met a barrage of protests from the scientific community. These subsided as it became clear that the principle would not be applied to fundamental research. Where the applications of Rothschild's ideas have broken down, however, has been in the failure of 'contracting' government departments to follow his prescription that they should also contribute to the long-term goals of the contracting research agencies.

Is history now repeating itself? The Advisory Council on Science and Technology (ACOST) has now suggested to William Waldegrave, the cabinet minister responsible for science, that the Rothschild principle be developed into a new 'purchaser-provider' basis for all 'mission-orientated'



research. In particular, this would mean formally separating funding agencies (such as the research councils) from institutions carrying out research. The Laboratory of Molecular Biology at Cambridge, for example, would no longer be a dependant of the Medical Research Council.

If adopted by the government in the white paper (policy document) on which Waldegrave is now working, the proposals would probably not make a major impact on top-quality research laboratories; they would be the likely winners in any open competition for funds. And pressures of market competition could well, as they have done with university departments, stimulate a more effective use of resources across the research community. But there are two dangers. First, ACOST's proposal may encourage the hegemony in government circles — meaning primarily the Treasury — of a way of thinking that marginalizes research for which no ready purchaser can be found. (This is the real danger of ACOST's parallel suggestion of a separate funding council for 'curiosity-driven' research.) Second, a range of research-related activities would be made more vulnerable to short-term political and economic pressures.

These are precisely the issues raised by the plan of the Department of Trade and Industry (DTI) to open up to public bidding the contract for managing geological data which all companies exploring for oil and gas in the North Sea are required to deposit with the government (see page 101). There is a clear economic logic in seeking the lowest bid (that is, the most cost-effective contractor) for this work. And there is no inherent reason why a private company should be any less diligent or competent than the British Geological Survey (BGS), which currently carries out the work.

But there are also wider considerations. For example, contracting out the work would inevitably undermine the overall effectiveness of the BGS, not to mention its ambition to develop a single national geological dataset, containing details obtained from both onshore and offshore prospecting. And while it may be safe to entrust responsibility for the offshore data to the DTI when it is in the interests of the British energy industry to do so, the department could also choose to dilute any long-term commitment if Britain's economic and industrial priorities change.

Most immediately, the new development has faced the BGS and its parent body, the Natural Environment Research Council, with a nasty dilemma: how to operate in the 'business-like' way required of a successful bidder for the contract (and, for that matter, in the manner now being demanded by the government of research council operations in general) while at the same time maintaining (and funding through claims for 'overheads') the science-base for which the research council has a statutory responsibility?

That highlights the central flaw in ACOST's proposals. Explicitly modelled (like Rothschild's original proposals) on the way large companies manage research, the purchaser-provider relationship might work if funds were sufficient to meet both short-term and long-term objectives. But, as the behaviour of such companies during the present recession has amply demonstrated, economic difficulties tend to lead to increased emphasis on the former, and to

reduce the importance attached to the latter. ACOST argues that sufficient checks and balances can be built into the purchaser-provider relationship to prevent this from happening. But, until British industry demonstrates a commitment to the long-term view in its own research spending, the argument that its approach to the organization and funding of science can be safely expanded into a national strategy should be regarded sceptically. □

Censorship ahead?

Dangerously, the British government is being impelled towards further restrictions on the British press.

THE old habit of executing the messenger when the news is bad still infects the British government, which is summoning up courage to impose legal restrictions on the British press. But ironically, and in a splendid illustration of how it habitually invites wild reporting of the kind it now proposes to prohibit by law, it has not made public the document that purportedly gives it a licence to act, but has allowed bits and pieces of it to be leaked. In the same spirit might a conductor threaten his orchestra with jail for performing badly after refusing to give them copies of the score.

The circumstances are these. For years, there has been discontent in the House of Commons and elsewhere at the intrusion into people's personal lives by the British press. Eighteen months ago, an inquiry under the lawyer Sir David Calcutt shook its head over some of the practices of which it had been told, decided that a law to ensure the privacy of private persons would be difficult to draft and concluded in an uneasy compromise that the government should give a statutory Press Complaints Commission a trial period to see what might be done by persuasion. Then, during a summer in which a cabinet minister lost office after reports of his affair with an actress, and after reports of the rocky marriage of the Prince and Princess of Wales were followed by news of their separation, Calcutt was asked to say whether he considered that the trial period had worked. He, it seems, has said NO.

But the difficulties that led to the first Calcutt compromise have not gone away. The leaking of his second document shows how, by letting information slip into public awareness in the manner it chooses, the government can control the temper of public discussion. Thus has public discussion of many important matters been engineered in recent years. The danger now is that there will be a stampede of opinion in favour of privacy, on the face of things an unexceptionable public good, and a definition of it that inevitably restricts the reporting of matters of importance further. Yet Britain already has legal procedures that allow prior restraint on publication, while journalists' failure to disclose their sources of information can be a crime. In the absence of positive declarations that information should be free, who can blame British publications for fearing that the censors are moving in? □

Bidding to manage North Sea data raises questions about privatisation

London. One important question raised by the British government's strategy for science is to what extent the private sector can be given responsibility for managing data on the country's natural resources — a task traditionally carried out by scientists at public institutions. This issue has surfaced after the Department of Trade and Industry (DTI), following new government guidelines, announced that it will shortly seek bids from private companies to manage geological data obtained from oil and gas exploration in the North Sea over the past 25 years.

The DTI says its approach will save money, but critics are concerned that it will undermine efforts to build up a national geological dataset and give the successful bidder an unfair commercial advantage. Companies engaged in North Sea exploration are now required to deposit geological data with what used to be the Department of Energy, which has been absorbed into DTI. In turn, the department has contracted out the management of this data through the Natural Environment Research Council to the British Geological Survey (BGS). Under a 1981 agreement, BGS's Hydrocarbons Unit in Edinburgh looks after both geological data and core samples from exploratory drilling submitted to the DTI — the latter in a storage facility built by the department — and provides the government with advice on the energy potential of its offshore areas.

BGS has been told that its contract to provide these services, worth about £2 million a year, will not be automatically renewed at the end of March. Instead, DTI officials will in the next few weeks put out contracts for both jobs, based on the principle that the cost of public services should be tested against those available on the open market.

DTI's decision is not thought to reflect criticism of BGS, and a department spokesman said last week that a bid from the BGS would be welcome. The department has already held preliminary discussions with several companies, and the new contracts are expected to be announced by the end of March.

Nevertheless, the decision has led BGS scientists to wonder whether a tendering process based primarily on cost could overlook some of the less tangible advantages of the current arrangement, in particular access to the survey's unique expertise. "I think it is outrageous that there should be a major [geological] data set over which the BGS may not have control", says Mike Dean, head of the hydrocarbons unit in Edinburgh.

Various industrial organizations — namely the UK Offshore Operators Asso-

ciation, UK Onshore Operators Group and the British Independent Explorers Group — are concerned that a successful bidder may face a possible conflict of interest if it also carries out consultancy work for individual companies. But the companies most likely to put in bids for the management contracts feel that such concern is exaggerated.

"If the BGS is providing a service to the DTI, why should other people not have the opportunity to provide the same service more efficiently?" says Simon Kendall of Simon Petroleum Technology in North Wales. "It is not against the national interest for the government to look for better value for money." Kendall says that it is "insulting" to suggest that the data would be less secure in private hands and points to his company's experience in handling confidential geological data for other countries.

Government officials are believed to feel

that the concerns of critics are misguided, since they acknowledge that the integrity of the North Sea databank is vital to the long-term prosperity of Britain's offshore oil industry. In contrast, the BGS claims that the oil-exploration industry supports its contention that the current arrangements should continue. In a recent submission to the government on the organization of science, it says that its good relationships with the oil industry could be jeopardized by the DTI's new policies on market testing.

"We do market testing like anyone else, but I am concerned about whether this is an appropriate application", says BGS director Peter Cook. "We have been curating data for 150 years. What we have here is a government department primarily concerned with policy, and any policy-oriented department is going to take a short-term view."

David Dickson

Cuenod will head Human Frontier programme

Tokyo. A Swiss neuroscientist will take charge of the Strasbourg secretariat of the Japanese-inspired international Human Frontier Science Program (HFSP), which supports research on the brain and biological functions.

The Strasbourg organization is expected to announce shortly that Michel Cuenod,



Michel Cuenod

director of the Brain Research Institute at the University of Zurich, has been chosen as secretary general. The appointment, welcomed in Japan, should reduce differences between Japanese founders of the programme and senior management in Strasbourg.

Relations between Japan, which provides most of the \$40 million a year spent by the programme, and the present secretary general, Sir James Gowans of Britain, have been strained since shortly after he took up his appointment in 1989. According to Japanese government officials and scientists, Gowans is out of tune with the basic philosophy of the programme; last year, Gowans decided to step down (see *Nature* 358, 527; 1992).

Cuenod was one of four candidates nominated by the nine members of the programme (Japan, the United States, Canada, France,

the United Kingdom, Italy, Germany, Switzerland and the European Communities). The others were Sydney Brenner of the University of Cambridge School of Medicine, Horst Sund of the University of Konstanz, Germany, and Lawrence Bogorad of Harvard University (see *Nature* 359, 567; 1992).

The field was reduced to Cuenod and Sund at a meeting in November of the HFSP council of scientists in Strasbourg. Later that month, at a meeting of the board of trustees with representatives of all members, Cuenod emerged as the favoured candidate. This Monday, Cuenod confirmed that he will accept an offer to take the job.

Cuenod now represents Switzerland in the HFSP science council and is popular with Japanese officials and scientists associated with the programme. At the November council meeting, Gowans and some other members of the council bitterly criticized Akiyoshi Wada, a former Japanese council member, for making public the differences of opinion between Japan and Gowans (see *Nature* 357, 356; 1992), while other members defended Wada. "It was like a shootout in a Western movie", according to one participant. Cuenod has voiced support for the basic philosophy of the programme as described by Wada.

Although Brenner was also popular with the Japanese, the British delegate did not support Brenner at the November council meeting (Brenner was nominated by Italy). That action effectively killed his candidacy.

David Swinbanks

French study sparks debate on informed consent laws

Paris. French researchers are calling for reform of laws on medical confidentiality, data protection and free and informed consent better to meet the needs of science following allegations in the press against a government research group.

The controversy involves a study by Christiane Capron, Michel Duyme and Michele Carlier in the laboratory of Pierre Roubertoux at the Centre National de la Recherche Scientifique (CNRS)/University of Paris V René Descartes Laboratory of Genetics, Neurogenetics and Behaviour in Paris. Building on work with adopted children (see *Nature* **340**, 552; 1989), the team assessed the pre- and post-natal effects of artificial insemination on the cognitive abilities of children with the same father to test whether assisted procreation itself could cause harmful neurological or psychological effects. Such effects have been found in mice.

Capron's group obtained from the operator of a private sperm bank, Sacha Geller, the names of 120 children born after artificial insemination. Geller is a controversial figure who created an agency for surrogate mothers (since banned) and who advocates artificial insemination as a way for single mothers and lesbians to have children. The CNRS group subsequently tested the cognitive abilities of the 120 children and 3,600 of their classmates.

L'Express magazine alleged in two stories last month that the researchers have contravened three laws — one forbidding doctors from telling anyone except another doctor about a patient, and then only if the information is used for therapeutic purposes; one protecting the confidentiality of data; and one requiring people involved in such matters to give their informed consent. The allegations have since been taken up by other media.

Claude Paoletti, director of life sciences at the CNRS, says that existing laws place researchers in a bind. At present, he says, a researcher's conscience and in-house ethics codes have been accepted as an adequate substitute for flaws in legislation; many scientists would find themselves breaking those laws, he maintains, if they were stringently enforced.

The former president of the national bioethics advisory committee, Jean Bernard, illustrates the dilemma facing researchers by citing the example of those who use epidemiological research to diagnose diseases with several causes but who, to do so, must first obtain the names of individuals. Paoletti says that similar problems arise whenever medical and basic research overlap and suggests that "there should be a law allowing researchers to access

medical information without breaking the law".

There may be progress on this front. A bioethics bill (see *Nature* **356**, 368; 1992) would for the first time allow medical information to be used for research with approval from a new committee. But the bill relates only to epidemiological research; scientists want legislation covering all relevant research.

The press reports also allege that the CNRS researchers did not have the free and informed consent of the children's parents. The letter sent to parents and teachers mentioned cognitive studies but not their precise aims. The requirement for such permission comes from a 1988 law designed for clinical trials, and scientists have questioned its relevance to research where the subject is not likely to be harmed.

Duyme says that the group did not seek the consent of parents because it would have "worried them unnecessarily". At the same time, he points out, the entire class was tested to preserve the anonymity of those being studied, following a protocol first approved in 1972 by the Ministry of Health.

Researchers perceive the controversy quite differently from the media. *L'Express*,



Carlier, Duyme and Roubertoux (from left) are at the center of controversy.

for example, portrayed the breach of confidentiality as a new and dangerous form of malpractice and its failure to obtain informed consent as deliberate subterfuge. But the author acknowledges that his perspective is coloured by his vociferous opposition to work on the genetic bases of human behaviour.

The controversy threatens to disrupt future research in the CNRS laboratory. In the meantime, French scientists are concerned that media review could be taking the place of peer review in judging the quality of their work.

Declan Butler

Medical council of Canada broadens role

Quebec. The role of the Medical Research Council (MRC) of Canada will be greatly expanded during the next five years, evolving from the country's major funding body for biomedical research to the "voice of the health sciences" in Canada.

In a strategic plan released this month, the MRC proposes to support research not just in traditional biomedical research areas such as basic biology, mechanisms of prevention, diagnosis and treatment and clinical trials but also in such subjects as the socioeconomic factors of illness, outcomes and cost-effectiveness of medical interventions and bioethics. The agency will seek partnerships with industry, in particular with the C\$6-billion (US\$4.5 billion) pharmaceutical industry (see *Nature* **359**, 351; 1992) and with those producing medical devices, computer software, electronics and chemicals.

MRC's new vision will require striking a balance between targeting research priorities and supporting investigator-initiated research. It will involve improved support for students, development of new classes of associates and awards and training programmes to increase the number of women

and minorities in the field. Current programmes will be evaluated, the peer review process reviewed and administration improved. Better procedures will be developed to communicate research results and to advertise the achievements of Canadian science. Attempts will be made to attract more funding for health research, including private support from organizations outside Canada.

The strategic plan grew out of what MRC president Henry Friesen has called "the most thorough self-appraisal in our history". That assessment was necessary, he said, because of the vast changes in recent years in biomedical research. The process involved more than 3,000 researchers and health administrators attending a series of meetings over five months that ended last May in their endorsement of an expanded role for the council.

In the end, most participants felt that the council had no choice but to expand its scope. As Mamoru Watanabe, dean of medicine at the University of Calgary, remarked: "If the MRC does not get involved [in these issues], it will not survive".

David Spurgeon

Protesters target European animal patents

Munich. A coalition of animal rights activists led by two British animal welfare groups is stepping up its campaign against animal patenting this week by formally opposing the granting by the European Patent Office (EPO) of a patent for the Harvard University Oncomouse. The opposition coincides with a pan-European campaign to prevent patenting of any animal or plant.

The British Union for Antivivisection and Compassion in World Farming are the forces behind a coalition of another 24 animal welfare groups in Europe. They believe that the granting of the patent last May contravenes a clause in the convention adopted in 1973 by the EPO's 16 signatories that prohibits patent for an invention whose exploitation may be "contrary to morality".

The opposition highlights the ambiguity of the convention's 'morality clause'. Written before the concept of transgenic animals was realized, it gives no helpful guidelines and leaves the definition of morality to the patent office.

The coalition disagrees with the EPO's interpretation, saying that engineering of animals designed to suffer is inherently immoral. It also points out that the patent covers the creation of transgenic animals never considered by the EPO.

In 1985, Harvard submitted an application for its Oncomouse to the EPO in Munich. Although a US patent was granted in 1988, the EPO threw out the application in July 1989, quoting the morality clause. But Harvard won on appeal after the EPO Board of Appeals decided that morality should be assessed "mainly on a careful weighing up of the suffering of animals and possible risks to the environment on the one hand, and the invention's usefulness to mankind on the other".

But the coalition also argues that the claim for usefulness is exaggerated, saying that animal models have limited value for human diseases. It also disputes the claim

that fewer animals will be used, saying that because of the high background of spontaneous cancers, the increased sensitivity of the Oncomouse to carcinogens will be offset by the high background of spontaneous tumours.

Objections to any patent may be submitted for up to nine months after it is granted, in this case by 13 February. The coalition expects to be joined by many other protest groups, including Germany's *Kein Patent auf Leben*, another umbrella group representing 48 individual pressure groups.

The EPO may take a year to respond. Although the principle that animals are patentable has now been established, individual applications (and there are now more than 100 pending) are considered on a case-by-case basis. The EPO also says that the recent granting of three further US patents on transgenic mice (see story below) will not affect its decision.

Coinciding with this growing protest is a debate within the parliament of the

European Communities (EC) on amendments to a proposed EC directive, whose aim is to give each member state equal patent protection for biotechnology inventions. The amendments, now in the hands of the European Commission, give some protection to animals but, more importantly, attempt to define the term morality. A move by the parliament to ban animal patents was narrowly defeated (by a vote of 96–104) last November. The final decision rests with the EC's Council of Ministers.

If the directive is passed with amendments that help to clarify the meaning of morality, it is likely that the 1973 law will be altered to conform to such wording. All EC countries, as well as several others, are signatories, and it is important that there is no conflict in the laws of the two bodies. Any reduction in the EPO's ability to address ethical issues is likely to ease tensions, even if the controversy over patenting of life forms is not laid to rest.

Alison Abbott

Ruling narrows US view of animal patents

San Francisco. A decision by the US Patent and Trademark Office to grant patents on three transgenic mice opens the gates to more patents on animals but suggests the government may be taking a more narrow view of the issue.

On 29 December, the agency issued patents on a Harvard-developed mouse model for prostate enlargement, known as 'Harvard II'; a mouse developed at Ohio University that produces human beta interferon, a natural protection against viral infections; and a mouse developed by GenPharm Inc. of Mountain View, California, that fails to develop mature T cells and would serve as a model for AIDS, rheumatoid arthritis and transplant rejection.

The only previous US patent on an animal was issued in 1988 for a mouse genetically engineered by Philip Leder and his colleagues at Harvard as a model for human cancers. That patent caused an uproar among animal-rights activists, environmentalists and some farmers, but a suit filed by animal activists against the patent office was not heard because the group could not demonstrate that any harm had been done. In addition, Congress has failed three times to enact a moratorium against patents on animals.

The length of time between the original "Harvard mouse" and last week's patents caused some to wonder whether the agency had adopted an unofficial moratorium. But Charles Warren, deputy director of the biotechnology patent examining group, says that the complexity of the claims was the

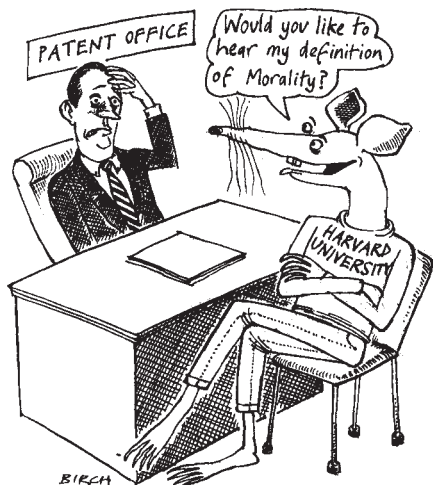
primary reason for the delay in issuing the patents. Ohio University filed its application in September 1987, followed by GenPharm in December 1988 and Harvard in February 1989.

Paul Clark, a lawyer with Fish & Richardson in Boston who prosecuted the Harvard II patent, said a major stumbling block had been a two-year argument over the breadth of the claim. Characterizing the agency's position as "retrenchment", Clark said government officials decided that genetic engineering on animals is unpredictable and that, therefore, a researcher could make claims only for those animal species on which the experiments had actually been done.

That policy differs from the agency's stance on the original Harvard mouse, which extends to alterations of the same gene in any mammal. Harvard II and the other new mouse patents apply to that species only. Future animal patents are likely to be similarly restricted, Clark predicted.

Warren said the patent office is considering 185 applications involving animals. At the agency's historical rate of allowance, only 50 or so are likely to be issued patents. Animals are being used as models for other human diseases as well as in research to improve the qualities of various farm animals. GenPharm, Genzyme Corporation of Cambridge, Massachusetts, and other companies are also genetically altering mammals such as mice, goats and cows to produce human proteins in their milk for therapeutic use.

Sally Lehrman



Czechs revise policy on research funding, aim for a more Western approach

Vienna. The newly created Czech Republic plans to develop a national policy towards research and to adopt a competitive system for awarding grants. That approach is in keeping with a report issued last month by the European Communities calling for university research to be given priority in the short term and for the government to focus on applied research and technological development to achieve long-term economic health. But the report warns that the changes must be implemented gradually to avoid damaging the country's fragile research infrastructure.

Science in Czechoslovakia was well-supported under communism, with levels of spending — at least 2 per cent of gross domestic product — comparable to Western Europe. But a large proportion of the money was wasted as the system had no single body responsible for an overall research strategy and gave out non-competitive awards. Most applied research was handled by government ministries, while basic research was the responsibility of the national Academy of Sciences.

Applied research institutes were run by

several different ministries and funded without regard to the quality or direction of the research; instead, political favouritism largely determined how the cake was divided. Chronically overstaffed, most of these

than 600 scientists but published on average only 12 papers a year. Even so, its funding has been reasonably well-protected since the 1989 revolution.

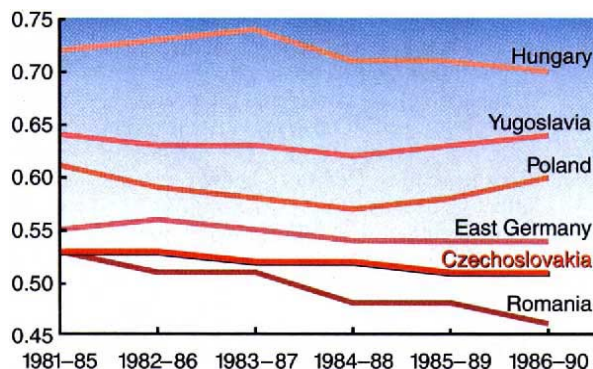
By contrast, industry-linked applied research institutes have fared badly. Government funding has dried up and industry, struggling to survive, has reduced its own commitment to the point where institutes are fighting over the few industrial contracts available. At least half the staff have lost their jobs.

Before 1989, the Academy of Sciences ran more than 70 institutes of varying quality. But because the money was given to the institute, rather than to a specific project, successful research did not attract extra funding. Universities received very little money for research, and only 14 of the republic's 24 institutes of higher education did any research.

Relations between the Academy of Sciences and universities have been strained since 1989. Resenting the academy's power under communism and their own lack of research facilities, the universities originally demanded the right to take over the research institutes. But the academy resisted,

East-bloc science: Impact of cited papers

(world average=1.0)



Source: Institute for Scientific Information, 1991

institutes had questionable relevance to the needs of their parent ministries and contributed little. One institute run by the agriculture ministry, for example, employed more

communism and their own lack of research facilities, the universities originally demanded the right to take over the research institutes. But the academy resisted,

Prague institute leads the way

Vienna. Ambitious plans are afoot for a centre of excellence, called the Prague Institute for Advanced Studies (PIAS), in the economically depressed capital of the new Czech Republic.



Peter Pechan

Born out of the idealism of the 1989 revolution, founded officially in 1991, and opened for operation this year, PIAS will be dedicated to research in the life sciences and to technology transfer.

The institute was originally linked to a sister organization in Slovakia, the Slovak Applied Research Centre (SARC), that would focus on the transfer of human skills for technological development. But the two institutes have gone separate ways since the decision, now in effect, to divide Czechoslovakia into two republics.

The future of SARC, which so far has a staff of only six, awaits larger decisions about the economy by the new government of Slovakia.

Prospects for PIAS are much brighter. Three of its five planned research divisions will be established by the end of the year and are already operating with a skeleton staff. The ecological medicine group has begun to assess the effects on health of air pollution in the mining and chemical areas around Teplice in northern Bohemia. Peter Pechan, the institute's

director, calls it "a model programme of cleanup to make up for our country's past". By 1995, divisions of biomedicine and imaging will join existing ones in plant biology and virology to complete the research teams.

A sixth division will organize technology transfer. PIAS has emphasized the development of intellectual property of commercial value, a new concept for the country, with researchers rewarded financially for their inventions. PIAS also wants to set up a science park that would benefit both basic research and industrial efforts.

This month, PIAS moves into a building formerly occupied by a government research institute. During the next five years, it expects to expand in size to more than 350 staff, all on short-term contracts, including collaborators from outside universities with which it intends to form close links. Its faculty will train as many as 50 PhD students a year.

But the best of intentions can be frustrated by harsh economic reality. PIAS will not come cheap, particularly since its dream is to become financially independent. PIAS has just received ECU490,000 (US\$650,000) from the European Communities to cover part of its 1993 expenses. It will receive an additional ECU10.9 million later in the year to cover the capital costs of equipping laboratories. Despite those contributions, the centre will need government contributions to meet an expected annual operating shortfall of as much as ECU4 million during its first five years.

A.A.

maintaining that the research enterprise would suffer if suddenly handed over to people without experience. To avoid a stalemate, the government turned for advice to the West.

The result was a large-scale study funded by a programme to help in the restructuring of Central European economies. The study, whose work was interrupted by the decision of the Czech Republic and Slovakia to go their separate ways, was conducted by the British consulting firm Segal Quince Wicksteed of Cambridge.

Some of the study's recommendations have already been implemented. Last month, for example, responsibility for science was transferred to the education ministry, which plans to establish a science unit to set national priorities. But the Ministry of Health will retain its own budget for medical research.

Efforts are also being made to create a system to make competitive grants. In 1991, part of the academy's budget was allocated on a competitive basis but the overall distribution of money did not change. The Ministry of Health introduced a system with a greater impact on funding decisions, and the Ministry of Education plans to use it as a model for an agency that would serve both universities and research institutes. The report advocates a 10-year evolution from a system giving nearly all the money directly to institutes to one allocating at least 85 per cent of its budget to individual projects.

Other recommendations will be implemented very soon. The report says that the first decade's grants should favour disadvantaged universities, giving them a chance to establish expertise in basic research, and it recommends that 20 per cent of the funds be set aside for collaborative work between universities and research institutes. But it argues against transferring the institutes to the higher education sector, saying the move would cause considerable disruption and might not foster such linkages. It wants centres of excellence dedicated to well-defined programmes of applied research, an idea in keeping with recommendations that the Czech Republic should focus on research encouraging long-term economic recovery.

The government is not expected to adopt all the recommendations. The report criticizes the government's decision to take a *laissez-faire* approach to technological development, reflecting a historical fear of authoritarian interference. An applied research centre in Prague (see left) is hoping to beat the odds. As usual, the major impediment to such an ambitious programme is money: since 1989, research funding has fallen by 50 per cent in real terms. In times of such economic hardship, the will to maintain a relatively expensive research base must remain strong if the system is to flourish.

Alison Abbott

Congressional pork cramps NASA's ground-based projects

Washington. Managers of the US National Aeronautics and Space Administration (NASA)'s Earth science programme are struggling to find money for hundreds of ground-based research projects scheduled for 1993. Unless money can be transferred within accounts, managers of NASA's \$150-million Earth sciences research and analysis programme — which includes fundamental studies unconnected to specific spacecraft missions — may, in the words of one researcher, "have to tell some investigators not to go into the field this year". The programme provides grants averaging \$100,000 a year for ground- and aircraft-based investigations.

The shortage, which could amount to as much as \$21 million, results in part from a decision by Congress to spend \$20 million this year on an Earth science information centre in Michigan, the home state of retiring US Representative Bob Traxler, who headed the appropriations subcommittee for NASA. The Consortium for International Earth Science Information Network (CIESIN) was not requested by NASA, nor was it peer-reviewed, making it one of hundreds of so-called "porkbarrel" projects. Traxler retired at the end of last year after serving 18 years in Congress.

To make room for CIESIN, Congress deleted \$15 million from the "mission operations and data analysis" portion of NASA's proposed budget in Earth science. That would have cut funds for analysing data from the Upper Atmosphere Research

Satellite (UARS) launched in 1991 primarily to investigate ozone depletion in the upper atmosphere. Last month, however, NASA and Congress agreed instead to take the money out of the research and analysis budget, temporarily rescuing UARS.

The research and analysis programme includes much of NASA's fundamental work in verifying data from remote-sensing observations in space, including flights over Antarctica and the Arctic to measure atmospheric ozone loss, as well as expeditions to study global tropospheric chemistry, solid Earth geophysics and ocean processes. The money is also used for laboratory research to calibrate the results of such observations.

NASA plans to eliminate some investigations scheduled for 1993 if the money is not restored, says UARS project director Robert McNeal, rather than applying the budget cuts across the board. "I'd rather do some things well than squeeze everybody to the point where nothing can be done", he says.

Berrien Moore, director of the University of New Hampshire's Institute for the Study of Earth, Oceans and Space and an adviser to NASA's Earth Observing System (EOS) project, says that this year's shortage demonstrates that the agency does not have the necessary resources both to build the spacecraft it has planned — including the multibillion-dollar EOS — and to analyse the results from those missions. A budget that's already constrained, Moore says, "can't possibly tolerate these congressional add-ons".

Tony Reichhardt

Chemical weapons treaty signed

London. After 20 years of formal diplomatic negotiation — and many more of lobbying from the scientific community — an international treaty banning the development, production and use of chemical weapons is being signed in Paris this week.

The treaty, likely to come into effect early in 1995, will prohibit all work on toxic chemical agents, apart from what is necessary for purely defensive purposes. It will also require all signatory states that currently possess such weapons — primarily the United States and Russia — to destroy them over the next 15 years.

Formally known as the Chemical Weapons Convention, the treaty is welcomed by scientists who see the technology as epitomizing society's misuse of modern science. "I think it is an astonishing success to have solved all of the many problems involved in a manner agreeable to a large number of states", says Matthew Meselson, professor of molecular biology at Harvard University.

Effective policing of the treaty will

depend largely on the operation of a new international agency, known as the Organization for the Prevention of Chemical Weapons (OPCW). Located in The Hague in the Netherlands, it is expected to take on a role comparable to that of the International Atomic Energy Agency in Vienna.

Several compromises have been necessary to reach final agreement. One involves the terms of 'challenge inspections', under which a signatory state can be required to demonstrate that fears of clandestine work on chemical weapons are unfounded. The United States — concerned that such challenges would be used to gather information from military facilities having little to do with chemicals production — has successfully argued that inspectors must wait for up to five days before entering a suspect site.

Similarly, the initial plan called for the destruction of existing stocks of chemical weapons within ten years. The United States and Russia are holding talks on appropriate technologies.

David Dickson



There is little room for crowds at Palmer Station in Antarctica.

NSF hangs out wary welcome sign

Palmer Station, Antarctica. US ambivalence towards tourism in Antarctica will be sharpened this month by the arrival of a ranger from the US National Parks Service. Her job will be to brief the tourists visiting this research station, the most accessible of the three on the continent.

The National Science Foundation (NSF), which manages the US Antarctic programme, is wary of tourists. Officials fear intrusion on scientific work and violations of the 31-year-old Antarctic Treaty, which technically prohibits even encounters with penguins that cause them to run away. But it is also anxious that those who do visit should form a good impression of Antarctica, whence the arrangement that the parks service will supply an "interpretative guide" for the current (Antarctic) summer season.

A dozen ships from various countries will carry 5,600 tourists this season to the west coast of the Antarctic Peninsula to admire its magnificent scenery. But only 1,200 will be allowed to visit Palmer, as passengers on the 12 cruises arranged by the four tour operators with which NSF has made prior arrangements. Their visits will be compressed into six weeks this month and next to avoid disruption to the many research projects that begin a new season in November and December. Until this year, either NSF personnel or employees from a US company hired by NSF to provide support services handled

all dealings with tourists.

Even so, visitors do not get a guided tour of Palmer but must satisfy their curiosity during an hour-long spell ashore to buy souvenirs at a portable gift shop and enjoy light refreshments in the dining hall. There they may rub shoulders with researchers, who appear at least as interesting to the tourists as the indigenous life-forms. A US tour operator says that the visitors most want to know "what it's like to live and work in Antarctica". Most scientists take these questions in their stride, believing that an educated public will help to preserve Antarctica as a research base.

Palmer's first ranger will be Maureen Loughlin, 36, a wildlife veterinarian with 14 years' service in several national parks. She says that her first task is to "see what works best for visitors"; upon returning to the United States she will develop whatever programmes, exhibits and publications are needed.

But her success may create more problems than it solves. NSF may then face pressure to make the position permanent, which would give an agency that does not have science as its chief mission a role in handling Antarctic affairs. Although it is unlikely that this season's pilot project will open the flood-gates to tourism, it suggests that the "Scientists Only" sign hung over the continent for so many years may not remain forever.

Jeffrey Mervis

US biotechnology groups merge

Washington. The US biotechnology industry has decided to speak with one voice on policy issues. This week, the Association for Biotechnology Companies (ABC) and the Industrial Biotechnology Association (IBA) announced plans to form the Biotechnology Industry Organization, bringing together IBA's 150 member companies with the 340 member firms of ABC. The new association will include more than 300 US biotechnology companies and represent more than 90 per cent of the industry's \$5.9 billion a year in product sales.

The two organizations have not always seen eye to eye. ABC and IBA took opposing positions over legislative efforts to trim the marketing exclusivity provisions of the Orphan Drug Act. ABC historically has represented smaller biotechnology companies.

In their first joint activity, the two associations asked President-elect Bill Clinton to support measures to stimulate long-term capital formation, strengthen process patent protection and allow the US Food and Drug Administration to collect user fees from companies.

Diane Gershon

NEWS IN BRIEF

Washington. Witnesses testifying before this year's science subcommittee in the US House of Representatives should expect to spend a good deal of time explaining themselves: only 6 of the 16 members of the subcommittee, whose members were chosen last week, have previously served in Congress, providing vivid evidence of the desire for change that voters expressed last autumn. The subcommittee itself has shrunk from 20 members, with a new subcommittee on technology, environment and aviation attracting the maximum allowable number, 33. US Representative George Brown (Democrat-California) remains chairman of the parent Science, Space and Technology Committee.

■ Ron Brown, the choice of President-elect Bill Clinton to be secretary of the Commerce Department, appears ready to give greater attention to environmental issues than his predecessors, who traditionally have focused on trade policy. In both oral testimony last week at his confirmation hearing before the US Senate and in written answers to questions, Brown talked about the significant role of the National Oceanic and Atmospheric Administration (NOAA) in monitoring the environment. But his plea for money to modernize the agency's fleet of ships reflects a favourite theme of the new administration: the investment will not only strengthen science, but it also will provide jobs for the US ship-building industry. Brown also strongly supported Clinton's campaign pledge to create more manufacturing technology centres as part of a more ambitious technology initiative within the Commerce Department. But he was careful to remind Congress that such programmes have yet to be drawn up.

■ Russia is now officially on board as a collaborator for the Superconducting Super Collider (SSC) being built in Texas after US Energy Secretary James Watkins last week signed an agreement between the two governments in the presence of Russian ambassador Vladimir Lukin. The US government regards the contracts with two Russian institutes to help design and build conventional magnets for two of the SSC's five accelerators (the low-energy and medium-energy boosters that will speed up opposing streams of protons on their way to the main ring, where superconducting magnets will help them reach energies of 20 GeV) as a way to lower the cost of the \$8.5-billion accelerator to US taxpayers. Roy Schwitters, director of the SSC laboratory, also points to the special talents of the Budker Institute for Nuclear Physics in Novosibirsk in designing and producing such magnets. The agreement will reduce by half the estimated \$200-million cost of the magnets.

Jeffrey Mervis

What lingua franca?

SIR — In your coverage of “Science in Japan”¹ you refer to the Japanese language as a barrier to “Japan’s full participation in international science”. I believe that this represents only a Western view. In terms of the total number of native speakers of a language, English comes a distant second to Chinese. In Asia alone, native speakers of Chinese live in China, Taiwan, Thailand, Hong-kong, Malaysia and Singapore. Most foreign students studying in Japan are also natives of these countries, as is also the case in Korea. Furthermore, the Japanese language boom is also seen predominantly in these countries. Because of the cultural relationships of the Chinese and Japanese languages, ethnic Chinese students find it easier to learn Japanese than English.

To identify a simplified universal language of science and communication, Margaret Mead noted that “there might be a written language to permit the visual representation of ideas, independent of existing languages, rather like the relationship of Chinese script with the various spoken languages of China”². How the Japanese language fits this description can be deduced as follows: (1) the Chinese use approximately 5,000 *kanji* ideograms; (2) the Japanese have reduced the *kanji* ideograms from 5,000 to nearly 2,000 for their daily use; (3) according to Daub *et al.*³, a foreigner can comprehend 90 per cent of the technical literature in Japanese if he or she can recognize 500–600 *kanji* ideograms.

For students whose native tongues are Chinese and Korean, this simplicity makes the Japanese language a useful tool to learn science.

Sachi Sri Kantha

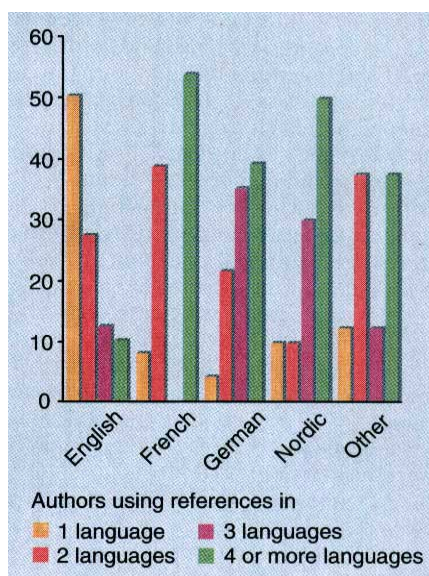
Department of Enzymes & Metabolism,
Osaka BioScience Institute,
6-2-4 Furuedai, Suita,
Osaka 565, Japan

1. *Nature* **359**, 573–582 (1992).

2. *New Scientist* 23 December 1971, 225.

3. Daub, E. E., Bird, R. B. & Inoue, N. *Comprehending Technical Japanese* (University of Tokyo Press, 1975).

SIR — There is a general trend to adopt English as a lingua franca in scientific communication. The director of the Library of Geological Sciences of the University of Barcelona, Mr Jordi Casadellà, and I are trying to analyse the actual situation of English as a lingua franca in geological sciences. To this end we have taken a sample of the recent geological journals and books to see how fast English is becoming the exclusive, or at least the main, language used in geological publications. In Western European countries in which the language of the people is not English, the



geological papers in reputable scientific journals are not predominantly English. At the same time, most English-speaking scientists tend almost exclusively to cite English-written scientific literature. This is a serious problem because it seems evident that a full appreciation of the breadth of geological knowledge makes necessary a knowledge of the scientific production of Western European countries, written mainly in French and German but also in Spanish and Italian and other languages. In geological sciences, the evidence that an author does not read languages other than English is a sign that his information is significantly incomplete.

Salvador Reguant

Universitat de Barcelona,
Facultat de Geologia,
Zona Universitaria de Pedralbes,
08071 Barcelona, Spain

Consciousness in other words

SIR — Despite the belief that linguistic confusion over the meaning of consciousness no longer forms a barrier to scientific investigation¹, further correspondence and reviews in *Nature*^{2–4} suggest that the definition of consciousness is still disturbing. I suggest that ‘consciousness’ is a term that is useful in everyday speech, but is confusing as a subject of serious study because the word is shorthand for a parcel of ideas.

One sense of the term ‘conscious’ is simple and uncontroversial: it is applicable to other individuals to mean that they are responding in a complex, non-reflex manner to stimuli. In this sense the word is used without embarrassment

of any animal whose reactions to stimuli we can interpret by comparison with our own; even etherized flies can be described as regaining consciousness. In this meaning, ‘consciousness’ can be replaced by ‘responsiveness’.

In its other meanings, a scientific approach is impeded by the perception that consciousness is detectable only as personal experience, and can be inferred in others only by acceptance of their reports of it. This being so, evidence for consciousness in others depends on human communication, and there is no way in which we can decide, except by supposition, whether the phenomenon exists in the absence of speech⁴. The persistence of the debate about its existence in animals also implies that we are unable to predict simple evidence for consciousness: it has no clear consequences for animal behaviour.

Even within this second type of meaning, I suggest that there are two ideas. One is that consciousness comprises the set of thoughts and sensations at any current moment; these are the internal counterpart of reactivity, and are by definition unavailable, except by report, to an outside investigator. Nevertheless, correlation of such reports with measurements of neural activity are a fruitful means of investigating the nature of sensation and ratiocination.

The second intrapersonal meaning of ‘consciousness’ is more complex; it is usually employed in the past tense and relates consciousness to memory³, for example in the remark “I was conscious of that event”. This statement can be replaced by “I can now remember that event”; conversely, if I say I was unconscious of an event, all I can legitimately state is that I cannot now recall it. Here consciousness is a projection into the past of the current experience of memory recall, giving a description of a supposed state of mind when the memory was laid down. It may be embarrassment at the fallibility of human memory that makes us prefer the more positive declaration “I was not then laying down memories” to “I cannot now remember”. Whether this is the case or not, it seems to me that this meaning of the word is essential to our understanding of the phenomenon of consciousness: since these are internal mental activities, only those that leave a record in memory have a continued existence and can become the basis of that experience which defines individuality.

John Clutterbuck

Department of Genetics,
Glasgow University,
Glasgow G11 5JS,

1. Gray, J. *Nature* **358**, 277 (1992).

2. Fletcher, H. L. *Nature* **359**, 665 (1992).

3. Longuet-Higgins, C. *Nature* **360**, 117 (1992).

4. Whiten, A. *Nature* **360**, 118 (1992).

Unreconstructed man

SIR — In the process of describing the physical characteristics of class I MHC molecules, which define the type of peptide that can be bound in their central groove, Peter Parham provides some sensational prose (*Nature* **360**, 300–301; 1992). What can he think of women?

He states that “class I molecules are promiscuous with regard to the peptides they bind . . . (yet are) . . . well adapted to any particular partner”. Then he gives us “port and cigars” and “fantasy” for gentlemen “before a sampling of their secrets” reveals “that the class I molecule is an intrinsically female structure” to resolve which “a cavalry of crystallographic gentlemen have charged”. It is most unfortunate, also, that the tee-hee fraternity refer to this structure as “Bjorkman’s groove” (she’s a woman, you know).

There is more in his piece but, best of all, with the complicity of *Nature*, there is a colour picture of a frankly vulval class I molecule filled by a pink penile peptide. If this was just foolish innuendo, I could forgive it as a good dose of tongue in cheek and hope that Parham grows up.

But Parham is wrong. Plugs/sockets and bolts/nuts are good examples of what engineers describe as male/female components in the only sense to which Parham could innocently refer. These components, however, have a low tolerance. They necessarily fit perfectly together in exactly the sense that MHC/peptide combinations cannot fit more than one peptide. Although if Parham really is using this analogy he manifests his most extreme chauvinism here by stating that class I grooves are promiscuous and suggesting peptides are not. The slag and the gentleman.

Looking deeper into Parham’s groove, I find the most insidious message in the implicit content of this puerile paper. This is a scientist writing about science; he should be painting pictures with an assembly of abstract facts. But his coital metaphor is a rich narrative. He provides a subjective story in which to couch observations — and to anticipate discovery. When do we get the orgasm of T-cell activation? And how many times can it happen? This is science for *Naturism* not *Nature*. It cannot be serious.

Harry White

University College London,
Department of Oncology,
91 Riding House Street,
London W1P 8BT, UK

SIR — I read the News and Views by Peter Parham with mixed feelings. As one studying cellular immune responses

during HIV infection, I was gladdened by the progress being made towards understanding the interaction of peptides with MHC class I molecules. On the other hand, I felt vaguely nauseated by the sexual innuendo throughout Parham’s commentary. To address this issue is to invite accusation that one is being oversensitive, a poor sport and driven by hormonal imbalances, but I shall take my chances. The picture of Parham and his colleagues sniggering over Van de Wall’s representation of the HLA-A2 molecule during seminars is disturbing enough. Benny Hill does immunology? To note that his text must have been approved by *Nature* is even more disquieting.

Charla Andrews

Aaron Diamond
AIDS Research Laboratory,
455 First Avenue, 7th Floor,
New York,
New York 10016, USA

SIR — I doubt if I am alone in finding the purple prose peppering Peter Parham’s article deeply offensive to the women who read this journal. Are we to understand that crystallography is a Victorian gentleman’s club where results are discussed over “port and cigars”? That the class I MHC “is intrinsically a female structure” (whatever that might mean), and that a “cavalry of crystallographic gentlemen have charged” to the task of resolving the shapes of bound peptides? For shame, Dr Parham, this type of prose is badly misplaced in a scientific community that includes women in its membership. No wonder many women scientists feel marginalized when senior colleagues persist in such sexist attitudes.

Lesley A. Bergmeier

Division of Immunology,
Guy’s and St Thomas’s Medical Schools,
London SE1 9RT, UK

Creative reviews

SIR — Michael Ignatius¹ gives convincing statistics to suggest that we have gone “review crazy”. As ‘Tristram Shandy’ wrote (1759): “Shall we for ever make new books, as apothecaries make new mixtures, by pouring only out of one vessel into another?” A review that simply collates the conclusions of all relevant papers in journalistic ‘bites’, sound or unsound, may be little improvement on the computer print-out from an abstracting service. Yet reviewing can be much more than an easy path to publications.

Well-written summaries, distinguish-

ing wood from trees, do help researchers, teachers and students. Reviews may also contain new ideas, cover papers in other languages, and show us which dead wood can safely be relegated to history. Some reviews, if that is the right word, draw on information not to be found in abstracts — from older work perhaps, or, to give a biological example, the many quantitative data on normal or control animals that are sometimes only incidental to the research.

As publication of ‘minimum publishable units’ accelerates, we need more creative reviews, not only to help us cope with the log-jam of new ideas but also to extract the most meaning from all those costly, hard-won data. Some editors discourage such synthesis in ‘experimental’ papers. Reviewing should not be denigrated: it is bad enough that it lacks the kudos of expensiveness.

Richard F. Burton

Institute of Physiology,
University of Glasgow,
Glasgow G12 8QQ, UK

Future of Baer Museum

SIR — I have just learnt that the Estonian Academy of Sciences intends to suppress the K. E. von Baer Museum, which was created in 1976 in Baer’s old house in Tartu, in order to rent the house to a private commercial organization for financial gain. Baer’s museum is now a centre for the history of biology, where Baer’s scientific papers and correspondence have been and are still collected and analysed (see *Nature* **353**, 104; 1991).

Ironically, the academy was the co-organizer, last February, of a successful international conference commemorating the 200th anniversary of Baer’s birth. Baer is rightly considered by the Estonians themselves to be their most famous scientist and his portrait has been chosen to illustrate the new Estonian Kr2 banknote.

The academy’s plan is of course influenced by the critical economic situation of the newly independent republic, but the damage it will cause to future Estonian and international researchers appears to be much greater than the immediate profit it will bring.

I hope concerned scientists will intercede with the president of the Estonian Academy of Sciences (Kohtu 6, EE 0106 Tallinn, Estonia) not to close Baer’s Museum.

Jean-Claude Beetschen

Centre de Biologie du Développement,
Université Paul-Sabatier,
31062 Toulouse,
France

Backwardness of human neuroanatomy

Francis Crick and Edward Jones

To interpret the activity of living human brains, their neuroanatomy must be known in detail. New techniques to do this are urgently needed, since most of the methods now used on monkeys cannot be used on humans.

OVER the past 20 years there have been great advances in understanding the neuroanatomy of the macaque monkey, especially its cerebral cortex. We have learned much about the functional parcellation of the monkey's cortex from both anatomical and physiological studies. We know, for example, that rather more than half the macaque's cortex is predominantly visual and that this visual part comprises more than 20 functional areas. A recent map¹ of these areas is shown in Fig. 1, and a provisional map of some of their interconnections is shown in Fig. 2. We also know something about the cortical layers from which many of these connections originate and to which layers they project. This knowledge, together with that showing the functional connectivity of other parts of the brain, will be essential as a foundation for the many neurophysiological experiments that will eventually show how the macaque brain functions.

In recent years there has also been remarkable progress in looking into the brains of living humans by means of magnetic resonance imaging (MRI)² and positron emission tomography (PET)³.

Most of the MRI scans used, although of high resolution, are static; they show structure but not activity. Such a scan can picture, for example, exactly how the cerebral cortex is folded in a particular individual but not what part is functionally active. The spatial resolution of classical MRI is now 1 mm or less so that the living brain can be seen in much better detail than before. But 1 mm³ of the human visual cortex may contain more than 40,000 neurons, so we are still looking at a low-resolution image.

A PET scan, although generally of lower resolution than MRI, shows where the brain is more active in particular tasks (or at rest) compared to a control task. By using an isotope of the oxygen of water, for example, the localized increase of capillary blood flow that accompanies local brain activity can be demonstrated. The present spatial resolution of most conventional PET scans is only 7–8 mm, but new photon detectors whose sensitivity approaches that of the theoretical spatial resolution of about 2 mm are being developed. The temporal resolution varies with the isotope used but is commonly minutes. Compare this with the temporal resolution of an electrode picking up the activity of a single neuron — of the order of 1 ms — a factor of 10⁵ or better. It is now possible to do both PET scans and an MRI scan on the same brain, so that the localization of activity can be made more precise.

Other types of scans exist, such as those that pick up minute variations of the magnetic field that accompany brain activity⁴. The MRI technique is being developed to make it amenable to the study of brain activity⁵. PET scanning is becoming increasingly used in conjunction with radioactive ligand binding to demonstrate changes in receptor densities in pathological and drug-dependent states⁶. All these scans are relatively expensive but they do provide information about the location of activity in the living human brain that can be obtained in no other way. To be useful their results must, in the long run, be related fairly precisely to the many different functional areas of the human brain. For this, detailed neuroanatomical data of the type shown in Fig. 1 are essential.

What is known about the neuroanatomy of the human brain? Do we have a human cortical map, corresponding to

that for the macaque shown in Fig. 1? And what does the human equivalent of the connectational map of Fig. 2 look like? The shameful answer is that we do not have such detailed maps because, for obvious reasons, most of the experimental methods used on the macaque brain cannot be used on humans.

It is possible to study, postmortem, the appearance of stained sections of the human cortex under the microscope and so observe differences between the basic structure of one region of the cortex and another, a subject called architectonics. It is particularly easy to locate certain cortical areas, for example, the first visual area (the striate cortex). Even with the naked eye one can see the

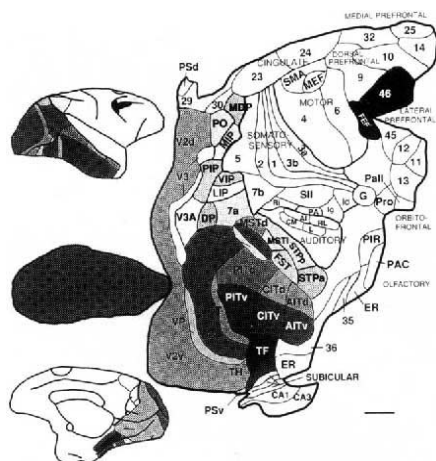


FIG. 1 Small drawings on the left, two views of the macaque cortex: upper, lateral view of the right hemisphere; lower, medial view. The main figure (different scale) shows the entire right cortical sheet unfolded and flattened out. To minimize distortion several cuts have been made round the edge of this sheet, such as the one outlining area V1 on the left. The shaded areas are all visual or partly visual. The main point of the figure is that there are many distinct cortical areas in the visual system of the macaque. Bar, 1 cm. (Copied, with permission, in black and white from ref. 1.)

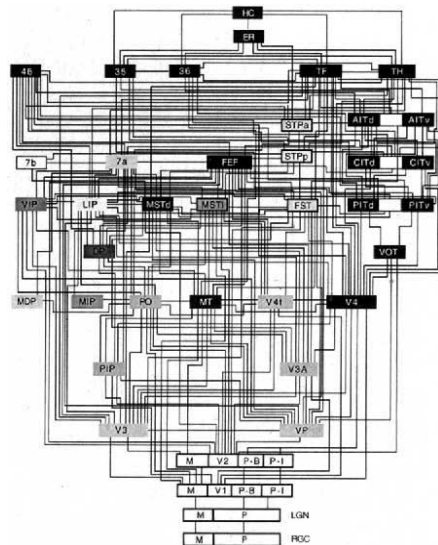


FIG. 2 Some of the intra-cortical connections of the main visual areas shown in Fig. 1. The abbreviations for the names of the cortical areas are (in most cases) the same in both drawings. The main flow of visual information from the eyes is from bottom to top. ER, entorhinal cortex; HC, hippocampus (marked CA1 and CA3 in Fig. 1); RGC, retinal ganglion cells; LGN, lateral geniculate nucleus of the thalamus. Above that is V1—the first visual area of the cortex, sometimes called the striate cortex or area 17. Each line between a pair of cortical areas denotes very many axonal connections, usually running in both directions. Connections to and from other cortical areas, especially involving areas near the top of the figure (such as frontal area 46) are not included. The main point is the great number of connections between the many distinct areas, though not every area is (strongly) connected to every other area. Modified by W. Suzuki and D. Amaral (personal communication) from ref. 1.

location of the striate cortex in a stained section of the human brain. In the early years of this century cytoarchitectonic methods enabled Brodmann to delimit at least 47 cortical areas including those he called 17 (the striate cortex), 18 and 19. We now know that, in the macaque, while area 17 is a single area, areas 18 and 19 have many finer subdivisions, as shown in Fig. 1. How then were the maps of Figs 1 and 2 obtained?

Several methods were used: for the visual areas a powerful physiological technique has been to study the response of single neurons to visual inputs. This is difficult to do in humans, if only because it requires opening the skull. Anatomically, the main method used in the macaque is to study the connections of small patches of cortex by injection of a substance that is transported along axons. After a delay of a few days, to allow sufficient active transport up or down the living axons, the animal is killed and the transported substance located microscopically in sections of the brain. A favourite tracer has been the enzyme horseradish peroxidase, succeeded recently by more sensitive tracers⁷. Using the anatomical method alone, it is possible to construct both Figs 1 and 2 for the macaque, though of course such techniques do not tell us how the neurons in each of these areas respond physiologically to visual or other inputs. Human neuroanatomy is so backward because we cannot use the common tracer methods on humans, since to transport the chemical the brain must be alive, and so the technique is ethically unacceptable.

We can provisionally make the assumption that the connectional map for the visual areas of the human cortex will be similar to that for the macaque, but this assumption will have to be checked. For other cortical regions, such as the language areas, we cannot use the macaque brain even as a rough guide as it probably lacks comparable regions.

It is intolerable that we do not have this information for the human brain. Without it there is little hope of understanding how our brains work except in the crudest way. The object of this paper is to make this need more widely known.

At this point we must make an important distinction between the cortical maps for a particular individual and those for the population as a whole. The macaque maps, shown in Figs 1 and 2, are average maps. The human cerebral cortex is not folded in exactly the same way in different individuals or even on the two sides of the brain in one individual, known from postmortem studies on human brains and from MRI reconstructions of living brains. In addition the exact position of some areas is known to differ somewhat from individual to individual. There are similar

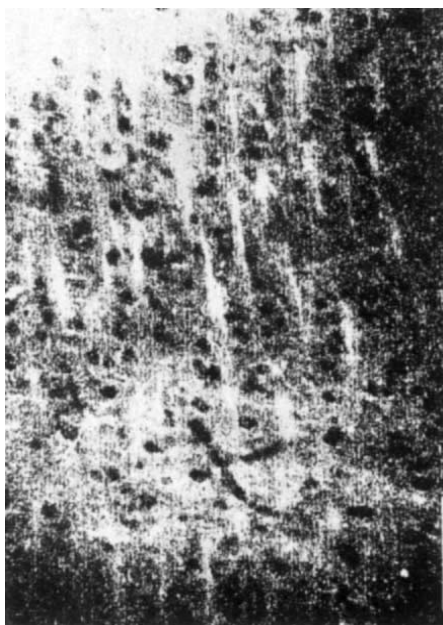


FIG. 3. Laser confocal image showing dil-labelled axons leaving an injection in superficial cortical layers (upper left), depositing terminal ramifications in layer VI and proceeding into the white matter. Injection made in auditory cortex of a monkey brain fixed with paraformaldehyde. Bar, 100 μ m.

differences in location between individual macaques (though they may be smaller). So far rather little variation has been found in the connectional map of the macaque between individuals but such variation has not been carefully looked for. A first aim, then, could be to construct the human equivalent of Figs 1 and 2 for an average human brain with the ultimate aim of discovering at least the location map (Fig. 1) for each individual whose brain is being studied.

There are several methods already in use on humans that could be improved and extended. For example Clark *et al.*⁸ have modified the MRI technique so that it is more sensitive to lipid and thus to myelinated axons. The degree of myelinization varies between cortical areas so that some degree of cortical localization should be possible and variations among individuals could be identified. New transcranial stimulation methods are also applicable to studying cortical localization and plasticity.

Various stains can be used on postmortem human brains, such as histochemical stains for the enzymes cytochrome oxidase and acetylcholinesterase. Immunocytochemistry and *in situ* hybridization histochemistry can be readily applied to localize many transmitter-related or structural proteins and their messenger RNAs, provided adequate fixation conditions can be assured. No doubt the human genome project will eventually throw up many more useful determinants of neuronal identity, though whether there will be a

different determinant for each distinct cortical area remains to be seen. Perhaps a combination of a few carefully chosen stains and labels would help considerably to establish homologies (or lack of them) between macaque brains and human brains. Such methods can, of course, only be used postmortem but methods for optimizing storage and recovery of human brain tissue are now available⁹.

Another new method that at last permits the tracing of connections in fixed postmortem material is the use of lipid stains such as the carbocyanine dye dil¹⁰ or one of its relatives. This spreads along axons by a diffusion process so that, in general, it is a slow method: to go 10 times as far takes 100 times as long. It could take many months to spread through the full extent of a long pathway, so there are time limitations on using it to establish the longer connections. Nevertheless, the method is now being optimized in experimental animals (Fig. 3) and with more work it should become applicable to fixed human brains¹¹.

Clearly what is needed for a modern human brain anatomy is the introduction of some radically new techniques, but unless there is a general awareness of the need for them they are not likely to arise. For such methods the brains from individuals who have just died may be needed. Although it is often not easy to get hold of fresh postmortem material, some effective brain banks have recently been established.

We wish we had more concrete suggestions for new techniques. Although we have not, we feel we should make a wide audience aware of this pressing need, especially as most neuroanatomists seem scarcely to have noticed it. If this article stimulates someone to devise suitable new methods to solve these problems, it will have served its purpose. $\square$

Francis Crick is at the Salk Institute, PO Box 85800, San Diego, California 92186; Edward Jones is in the Department of Anatomy and Neurobiology at the University of California, Irvine, California 92715, USA.

1. Felleman, D.J. & Van Essen, D. *Cerebral Cortex* **1**, 1–48 (1991).
2. Edelman, R.R. *Disc. Neurosci.* **7**, 11–63 (1990).
3. Raichle, M.E. *Cold Spring Harb. Symp. quant. Biol.* **55**, 983–986 (1990).
4. Hari, R. & Lounasmaa, O.V. *Science* **244**, 432–436 (1989).
5. Tank, D.W., Ogawa, S. & Ugurbil, K. *Curr. Biol.* **2**, 525–528 (1992).
6. Wong, D. F. *et al. Science* **234**, 1558–1563 (1986).
7. Heimer, L. & Zaborsky, L. (eds) *Neuroanatomical Tract-Tracing Methods 2* (Plenum, New York, 1989).
8. Clark, V. P., Courchesne, E. & Grafe, M. *Cerebral Cortex* **2**, 417–424 (1992).
9. Jones, E. G., Henry, S.H.C., Liu, X.-B., Potkin, S.G. & Tourtellotte, W.W.J. *Neurosci. Meth.* **44**, 133–144 (1992).
10. Godement, P., Vanselow, J., Thanos, S. & Bonhoeffer, F. *Development* **101**, 697–713 (1987).
11. Burkhalter, A. & Bernardo, K.L. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1071–1075 (1989).

Dark matter comes in from the cold

The year started with a bang for astronomers and cosmologists gathered at Phoenix, Arizona, last week, with the report of observations indicating that the Universe is made largely of exotic dark material.

HUMBLE baryons, the nuclear material from which our Solar System is made, represent a minuscule fraction, about 4 per cent, of the matter in the Universe. Most of the matter is dark, detectable by its gravitational effects alone. Thus runs the recurrent theme of the Big Bang cosmology to which most astrophysicists subscribe. Observational astronomers, however, urge caution, for solid evidence has remained elusive.

But 1993 has begun with release of new observations that promise to resolve some aspects of the controversies still endemic to the Big Bang cosmology. For the first time, diffuse hot intergalactic gas has been detected, by the ROSAT X-ray satellite, in a small group of galaxies (J. S. Mulchaey *et al.* *Astrophys. J. Lett.*, in the press). As described by Mulchaey in Phoenix last week*, the gas distribution is centred on the group, and the gas temperature is about 10 million degrees kelvin. Knowledge of the gas temperature with the reasonable assumption that the gas is gravitationally confined allows an estimate of the total mass of the group. Baryons, seen as hot gas and in the luminous galaxies, amount to about 4 per cent (and at most 15 per cent) of the total mass. The remainder is dark matter. For the first time, one can conclusively state that diffuse dark matter dominates the gravitational potential of galaxy groups on megaparsec scales. The more dramatic consequence, however, is that we can now aspire to ascertain the nature of the dark matter.

ROSAT's X-ray scan of the NGC2300 group gives not only the gas temperature but also the abundance of heavy elements, formed in stars long after the Big Bang. The spectrum does not have enough resolution to detect individual elements, but taking the element ratios to be solar, Mulchaey and colleagues find the abundance of elements heavier than lithium to be about 6 per cent (and no more than 20 per cent) of their abundance in the Sun. Therefore the gas must be predominantly primordial and predates the galaxies.

Nucleosynthesis in the first few minutes of the Big Bang provides a remarkably concordant explanation of the abundances of the light elements: helium, deuterium and lithium. Theory

and observation coincide only if the baryon density in the Universe lies within a narrow range: between 4 and 8 per cent of the dark-matter density in a universe at critical density. New evidence that the Universe is indeed at a critical density, destined to expand forever, but at ever decreasing speed, is presented by K. Kellermann elsewhere in this issue (see the News and Views article by P. Scheuer, overleaf).

Hitherto, astronomers have been able to compare gas and dark matter over large scales only in rich galaxy clusters, where they find diffuse intergalactic gas amounting to at least 10 and even 30 per cent of the dark-matter density in some cases. But that gas is far from pristine, being highly contaminated by ejecta from galaxies — astronomers eagerly await the scheduled launch of the Japanese X-ray satellite ASTRO-D next month to help ascertain the precise location within the clusters of the enriched gas component. Moreover, the complex dynamics of gravitational collapse may operate differentially on gas and dark matter: the X-ray surface brightness and the clumpy distribution of some rich clusters point to this, as does the mapping of dark matter by gravitational lensing. The gas content of rich clusters, even on their outskirts, need not be a good measure of the initial ratio of baryons to dark matter. Only groups of galaxies like NGC2300 can do the trick.

Could the nucleosynthesis prediction be in error? There is of course the cosmological deduction that only a small number (less than four) types of neutrinos are allowed for a unique and simple fit to observations of light-element abundances. The LEP experiments at CERN, as well as results from the Stanford Linear Collider, converged on precisely three types of neutrinos from studies of neutral Z-boson decays. The prediction also relies on the Hubble constant, measuring the expansion of the Universe, and on the temperature of the cosmic microwave background. Only a Hubble constant of about $50 \text{ km s}^{-1} \text{ Mpc}^{-1}$ yields an age for the Universe that is consistent with a universe at the critical density and with the ages of the oldest stars. The direct measures of the Hubble constant (involving supernovae, gravitational lensing, microwave background 'absorption' towards galaxy clusters) all favour such a low value,

although indirect techniques mostly indicate a somewhat higher value.

The cosmic microwave background is the remaining ingredient. Its temperature has now been measured to be that of a blackbody with deviations no larger than two-hundredths of one per cent near the peak intensity at a wavelength of 1 mm. The temperature is measured to be 2.726 K, with an uncertainty of only 0.01 K, as reported at Phoenix by John Mather, project scientist for the COBE satellite. Such precision in a cosmological measurement is unprecedented. Combination of the microwave background temperature, the Hubble constant and the observations of the light-element abundances leads to an unambiguous prediction of the baryon fraction. Its measurement in diffuse hot gas that is largely uncontaminated by gas stripped from galaxies points towards a Universe at critical density.

The dark matter is the stuff of particle physicist's dreams, some yet-to-be-discovered form of nonbaryonic matter. Examples abound: from axions, weighing in at 10^{-14} proton masses to supersymmetric relics such as photinos of 100 proton masses or more. We may have to wait until the next millennium for the Superconducting Super Collider and CERN's Large Hadron Collider to suggest anything more definitive. Or perhaps the particles, if they constitute some large fraction of the dark matter of our own galactic halo, will be detected directly.

Nor need seekers of baryonic dark matter be unhappy: the mass detected in luminous baryonic matter — stars and gas — amounts to at most one per cent of the critical cosmological density. The confirmation of a baryon-to-dark-matter fraction of a few per cent, which must await further ROSAT mapping of other galaxy groups, would practically guarantee that the dark halo of our Galaxy is teeming with compact baryonic fragments, most probably aborted stars, brown dwarfs or stellar remnants (old, cold white dwarfs). We can be certain that dark matter is out there, in amounts on the largest scales that totally dominate the dim glimmer of distant galaxies.

Joseph Silk

Joseph Silk is in the Departments of Astronomy and Physics, University of California, Berkeley, California 94720, USA.

*American Astronomical Society, 181st National Meeting, Phoenix, 3–7 January 1993.

Weighing the Universe

Peter Scheuer

THE Universe expands, but its expansion is continually slowed down by the gravitational attraction between its constituent bits, and it has always been an open question whether the bits will ultimately escape from each other (as a sufficiently fast rocket escapes from Earth) or whether they will collapse in a heap. On page 134 of this issue¹, Kellermann presents another piece of evidence indicating that the Universe may be close to the dividing line between these two possibilities.

The strength of gravitational forces relative to the momentum of the expansion is measured by a dimensionless density parameter Ω or a deceleration parameter q_0 , and so long as one abstains from introducing a 'cosmological constant' (a mystic repulsion that increases with distance) they are related simply by $\Omega=2q_0$. Models with $\Omega \leq 1$ are open and expand forever; models with $\Omega > 1$ are closed and doomed to collapse. Heroic efforts were made to measure q_0 by measuring the brightness (magnitude) and redshift (distance or age) of distant galaxies and thereby tracing the geometrical history of the Universe, but ultimately they met with little success, partly because of the difficulty of observing faint objects, but more fundamentally because one could not disentangle geometry from the evolution of galaxies with epoch. That is why, in recent times, the most plausible evidence on Ω has come from very indirect arguments.

Coincidence

Theorists have long had reason to wish for $\Omega=1$. The point is that, once gravity begins to win, it wins more and more strongly; conversely, once expansion wins, gravity becomes less and less significant. If Ω is (say) 0.5 now, it must have been 0.9 at redshift $z=10$ (when the Universe was a small fraction of its current age), 0.999 at $z=1,000$, and so on. Thus Ω had to start extremely close to unity in the early Universe, and it is hard to believe such an extraordinary coincidence unless there is some good reason (even if we don't understand what it is) why Ω should be exactly 1. Sceptics might dismiss that as mere aesthetic preference, but in the past decade, inflationary theories of the early Universe have also predicted that $\Omega=1$ on the basis of physics, albeit speculative physics.

The amount of matter that astronomers can actually observe in the form of luminous stars and gas is at most a few per cent of the closure density, $\Omega=1$,

and some recent estimates² are below 0.5 per cent. It is not enough to bind clusters of galaxies or even to account for the rotational velocities in the outer parts of spiral galaxies. Some of the missing mass could be in the form of brown dwarfs and other underluminous stuff; Arp pointed out long ago that a closure density of scientific journals would be unobservable, and fully ionized intergalactic gas is also pretty elusive. However, nucleosynthesis in the first few minutes of the Universe limits the total density in ordinary matter quite severely: to create the observed abundances of D, ³He and ⁷Li (which are destroyed rather than created in stars) requires $\Omega_{\text{baryons}}=(0.06 \pm 0.02)h_{50}^{-2}$, where h_{50} is the Hubble constant in units of 50 km s⁻¹ Mpc⁻¹ (ref. 3). If there is more dark matter, it must be neutrinos with a non-zero mass, or axions, or some other exotic stuff.

There are lines of argument based on observation which favour $\Omega \approx 1$; they depend on the growth of irregularities. Determined campaigns of redshift measurements have revealed a local departure from a uniform Hubble expansion, and one would like to attribute these peculiar velocities to the gravitational pull of local mass concentrations revealed by the measured distribution of galaxies. This work leads the astronomers involved to $\Omega_{\text{total}}^{0.6}b \approx 1$ where $b=1$ if one assumes that light traces mass, and b is a 'bias parameter' if one assumes, slightly less optimistically, that galaxies are b times more strongly concentrated than total mass. It seems implausible that b is much less than 1, so the implication is that $\Omega \approx 1$. Contrariwise, Efstathiou *et al.*⁴ find evidence for $\Omega \approx 0.2$ in the form of the angular correlation function of the distribution of galaxies; they discuss a range of fluctuation measurements, including the COBE (Cosmic Background Explorer) estimates of fluctuations in the microwave background, and find that no combination of parameters gives a satisfactory account of all the data.

The discovery of gravitationally lensed quasars revived hopes (as yet unfulfilled) of a direct geometric determination of the Hubble expansion constant H_0 and perhaps even of q_0 . Kellermann also uses a direct geometric approach, but of a different kind. As he reminds us in his article, the principle of his method was advocated by Hoyle way back in 1958⁵. Hoyle found, with evident surprise, that, in the standard Big Bang cosmologies, the angular size of a distant 'standard rod' goes through a minimum at $z \approx 1$

and increases again at larger z . He wrote "A remarkable difference between the theories now emerges. . . . The asymptotic value [of the angular size of a radio galaxy like Cygnus A] can never be less than 15" arc for E-dS [Einstein-de Sitter] but can fall as low as 4" arc for SS [steady-state cosmology]". We need not be too surprised by Hoyle's mathematical result if we use the (only slightly specious) argument that the Universe was really very small in the old days, so a standard rod at high z was quite close to us when it emitted the light that we see now.

Surveys

Fifteen years later, adequate maps of radio galaxies and quasars had become available in large numbers, and the data agreed with none of the cosmological models; the angular sizes of the largest radio sources went on decreasing as $1/z$ even at large redshifts z , just as they would in euclidean space. Once again, it seemed to cosmologists (reluctant to jettison all their beliefs) that evolutionary changes in the population of supposedly standard objects overwhelms the geometrical effects. (It has also been suggested that we observe mostly very powerful sources at large z and that the correlation is between size and power, not size and epoch.)

The new feature of Kellermann's analysis is that he uses the sizes of the compact jets in the nuclei of quasars (and of many radio galaxies too), which are tens of hundreds of light-years long, rather than the total length of the radio source which is typically 10^5 – 10^6 light-years. He argues that these compact structures are well within the host galaxy and short-lived, so they are well sheltered from external influences and should develop in much the same way at all epochs. The price to be paid for this improvement is that, while the large-scale structure of a radio source is sharply bounded, the small-scale jets just fade with distance from the nucleus, so one has to make somewhat arbitrary rules defining their lengths. That price is perhaps not as great as it first seems, because the scatter in intrinsic sizes is great in both large- and small-scale structures; furthermore, Kellermann points out that the strongest selection effect in his data tends to diminish the derived value of Ω , so that his result of $\Omega \approx 1$ is, if anything, an underestimate.

1. Kellermann, K. I. *Nature* **361**, 134–136 (1993).
2. Persic, M. & Salucci, P. *Mon. Not. R. astr. Soc.* **258**, 14P–18P (1992).
3. Peebles, P. J. E., Schramm, D. N., Turner, E. L. & Kron, R. G. *Nature* **352**, 769–776 (1991).
4. Efstathiou, G., Bond, J. R. & White, S. D. M. *Mon. Not. R. astr. Soc.* **258**, 1P–6P (1992).
5. Hoyle, F. in *Paris Symp. Radio Astronomy, IAU Symp.* No. 9 (ed. Bracewell, R.) 529–532 (1959).
6. Peacock, J. A. *Mon. Not. R. astr. Soc.* **223**, 113 (1986).

Assuming that Kellermann's analysis really turns out to tell us about Ω (and not some subtle, indirect evolutionary change in the way active galactic nuclei operate), what Ω does it measure? It does not discriminate between different sorts of mass, whether ordinary or exotic, but it does depend on the distribution of that mass in space. Objects seen through low-density regions between galaxies would suffer the converse of gravitational lensing, and objects seen through mass concentrations will be lensed and will generally appear bigger

and brighter. Indeed, astronomers have worried that selection in favour of the lensed, and thus apparently brighter, objects could seriously affect, for example, counts of radio sources. Detailed estimates⁶ indicate that such effects are weak in the case of source counts; analogous sums for Kellermann's analysis will no doubt keep a student somewhere out of mischief. □

Peter Scheuer is at the Mullard Radio Astronomy Observatory, Cavendish Laboratory, Cambridge CB3 0HE, UK.

VISUAL SYSTEM

Vision in blind mole rats

Jon H. Kaas

THE term 'blindsight' was invented to acknowledge the rather paradoxical observation that humans with extensive damage to the visual cortex and no awareness of their own visual capabilities can nevertheless perform a number of visually guided tasks¹. This is a dramatic indication that the visual 'system' is not a single processing stream with a single function, such as object identification, but a combination of subsystems, each depending on light that reaches the retina as a source of information, but then using the information in different ways and depending on different brain structures. Cooper *et al.*, in a report on page 156 of this issue² on the structural organization of the visual system in behaviourally blind mole rats, demonstrate in an equally dramatic fashion the parcelling of the visual system into subsystems with quite different functions.

For most investigators, mole rats would hardly be thought of as suitable subjects for studies of the visual system. These mole-like rodents live in underground tunnels and would seem to have little use for vision. Indeed, they are distinguished from all other rodents by the absence of an opening in the skin for the eyes and by having the smallest eyes of any mammal. Cooper *et al.* report that the two major central targets of the eye, the lateral geniculate nucleus and the superior colliculus, are so greatly reduced in size in the mole rat that they are almost absent. This is what we would expect from other comparative and experimental studies. In true moles, for example, the eye is reduced in size and the central targets of the eye are also reduced³. Furthermore, because the development of different levels of the visual system are linked, any reduction or loss of the eye should be followed by reductions in central visual structures. Thus, rodents with mutations that lead to the reduction or absence of eyes⁴, and

rodents blinded early in development⁵, have small superior colliculi and lateral geniculate nuclei.

In experiments of considerable technical difficulty, Cooper *et al.* managed to inject tracers into the rudimentary eyes of mole rats and demonstrate that most, but not all, of the central targets of the retina are greatly reduced. In contrast to the diminished pathways to the lateral geniculate nucleus, normally used for perception and object identification, and the diminished pathways to the superior colliculus, normally used for object localization, the projections to the hypothalamus are not only preserved, but are even enhanced relative to those of other rodents.

In mammals, the retinal inputs to the suprachiasmatic nuclei of the hypothalamus provide information about day lengths and thereby regulate biological rhythms or cycles that are important in reproduction and other daily and seasonal activity patterns⁶. Apparently, the tiny eyes beneath the skin of mole rats living in dark tunnels receive enough light to provide useful information about daily changes in ambient light levels, and nearly all of their remaining visual system is relegated to processing this information. Information on light levels and durations may be critical in timing the mating season for mole rats, and in regulating patterns of feeding and thermoregulatory behaviour^{2,6}.

The selective combination of reduction with preservation or enhancement of different components of the visual system of mole rats constitutes an instructive example of mosaic evolution — the concept that different components or characters of complex systems can change independently⁷. How this was achieved in mole rats is unknown, but the heterogeneous pattern of alterations argues against the common supposition that the loss of the visual system in

mole rats and other animals adapted to reduced visual environments is simply a consequence of relaxed selection pressure when vision is unused. Cooper *et al.* point out that there is a metabolic cost involved in maintaining a large visual system, and this cost can be especially significant in animals with a limited food supply. Thus, mole rats adaptively reduce behaviourally irrelevant components of the visual system while maintaining and enlarging components that remain important.

Mole rats also illustrate a solution to the general problem of scaling in biological systems. One of the basic interests of the eminent comparative physiologist Knut Schmidt-Nielsen has been how animals and parts of animals are structurally modified in order to preserve function when they become larger or smaller⁸. Reducing the size of the eye in mole rats or any other vertebrate would introduce several problems related to visual acuity and perception⁹. Most obviously, a very small eye would form a very small image on the retina, and without a corresponding but impossible reduction in the sizes of the photoreceptors, image quality would greatly suffer.

Furthermore, the ganglion cells and other neurons that process information from receptors and relay it to more central brain structures cannot be appreciably reduced in size. Thus, in very small eyes, photoreceptors and retinal ganglion cells are greatly reduced in number. Object identification and localization depend on information from large numbers of different classes of receptors and ganglion cells, although the numbers and classes of receptors and ganglion cells vary with the degree of precision needed for these tasks and the extent to which colour discrimination is included¹⁰. Mole rats sidestep most of these problems of scaling down the size of the eye by abandoning functions that are most impaired by size reduction, and devoting the meagre remnant of preserved receptors and ganglion cells to light detection and possibly to sex and well-being. □

Jon H. Kaas is in the Department of Psychology, Vanderbilt University, Nashville, Tennessee 37420, USA.

1. Weiskrantz, L. *Blindsight* (Clarendon, Oxford, 1986).

2. Cooper, H. M., Herbin, M. & Nevo, E. *Nature* **361**, 156–159 (1993).

3. Lund, R. A. & Lund, J. S. *Exp. Neurol.* **13**, 302–316 (1965).

4. Cullen, M. J. & Kaiserman-Abvamos, I. *J. Neurocytol.* **5**, 407–424 (1976).

5. DeLong, G. R. & Sidman, R. L. *J. comp. Neurol.* **118**, 205–219 (1962).

6. Turek, F. W. *A. Rev. Physiol.* **47**, 49–64 (1985).

7. Futuyma, D. J. *Evolutionary Biology* (Sinauer, Sunderland, 1986).

8. Schmidt-Nielsen, K. *Scaling: Why is Animal Size so Important?* (Cambridge University Press, 1984).

9. Barlow, H. B. *Proc. R. Soc. Lond.* **B212**, 1–34 (1981).

10. Rodieck, R. W. *The Vertebrate Retina* (Freeman, San Francisco, 1973).

Keeping that youthful look

Sean C. Solomon

THE record of impact cratering on Venus, recently revealed from high-resolution radar imaging by the Magellan spacecraft, has been interpreted as indicating that the planet underwent catastrophic global resurfacing about 500 million years ago¹. Although other interpretations of the crater characteristics have been suggested², the possibility of geologically rapid global resurfacing on the planet that most resembles our own in terms of mass, density and bulk composition has generated widespread interest: what mechanisms could cause such a catastrophe? Several ideas were discussed at the Fall Meeting of the American Geophysical Union last month*.

The average age of Venus's surface has been estimated from the number of larger impact craters (the planet's thick atmosphere has tended to protect it from the smaller bodies that would give craters smaller than 30 km in diameter) and from the cratering rate on Earth and the Moon, or from estimates of the number of asteroids whose orbits cross that of Venus³. The answer, 500 million years^{1,2} (a tenth of the age of the Solar System), is not in itself remarkable for a planet which, like the Earth, has a hot, dynamic interior — despite the absence of significant erosion. What is remarkable is that the craters at all sizes are indistinguishable from a random population (that is, spatial variations in crater density are not larger than those due to chance) and almost none has been significantly modified by tectonic strain or by volcanic flows beyond the crater rim^{1,2} (Fig. 1), even though Magellan images show volcanic and tectonic features to be nearly ubiquitous on Venus.

The interpretation of Schaber and

others¹ is that most of the surface dates from the end of a global resurfacing event that ceased about 500 million years ago, and that the small fraction of craters volcanically embayed or modified by deformation indicate that volcanic and tectonic activity has since been at much lower levels. Phillips and colleagues² have argued that a correlation of locations of modified craters with areas of low crater density and an inverse correlation between crater density and Magellan radar backscatter (a quantity elevated in regions of high topography and high roughness, both thought to be signatures of comparative geological youth) indicate that the Venus surface exhibits a spectrum of ages. (The paucity of small craters, however, hampers the use of crater density in determining the relative ages of geological units, a method used with the other solid planets and satellites.) Phillips *et al.* have described an alternative model in which resurfacing occurs episodically in small patches a few hundred kilometres in extent, with a characteristic time between events of the order of 10^5 years (Fig. 2).

Nonetheless, the catastrophic hypothesis and its possible mechanisms have captured the lion's share of the attention. At the meeting, D. L. Turcotte (Cornell University) proposed that global lithospheric overturn operates episodically on Venus, and that for the past 500 million years the lithosphere has been cooling and mechanically stable. Although the mechanism for stabilization of the lithosphere is not an explicit component of Turcotte's scheme, he argues that a cool and thick lithosphere is in better agreement (than one in steady-state conductive equilibrium) with the large values obtained from topographic profiles for the rigidity of the lithosphere at sites of presumed lithospheric bending⁴ and with the large ratio of long-wavelength gravity anomaly to long-wavelength topographic relief at highland areas⁵.

Parmentier and Hess⁶, following Stevenson and Bittker⁷, have recently suggested that lithospheric stabilization may occur on Venus because of a decrease in the density of lithospheric mantle following melt extraction. But they add that the cooling

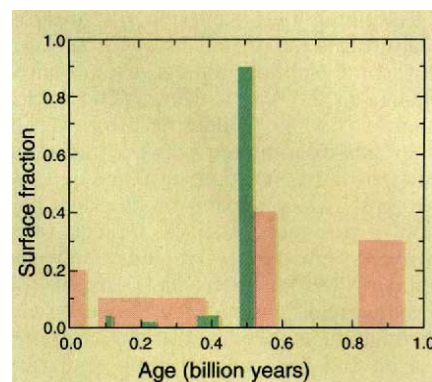


FIG. 2 Two schematic distributions of surface age on Venus. With catastrophic resurfacing¹ (green histogram), most of the surface is about 500 million years old, but small areas have younger ages. With episodic resurfacing² (red histogram), there is a more widespread distribution of surface ages from 0 to approximately 1,000 million years. Both possibilities are broadly consistent with the distribution and states of preservation of impact craters on Venus. (Courtesy of R. J. Phillips.)

of such a layer may subsequently raise its density to the point where it becomes unstable. Global overturn of this unstable layer, which would sink back into the mantle, would be followed by partial melting of the upper mantle, global resurfacing, and the gradual development of a new buoyant layer and another extended period of stability. In simple one-dimensional models of this process, the time between lithospheric instability events is 300–500 million years.

Others point to time-variable mantle convection rather than lithospheric instability as a mechanism for global resurfacing. An early effort to simulate such convection on Venus⁸ led to models in which the characteristic flow speed and mantle heat flux showed large variations (by factors of 2–10) at intervals of 100–200 million years. J. Arkani-Hamed (McGill University) reported improved calculations with better spatial resolution which are still time-varying but which show significantly weaker fluctuations of the order of several tens of a per cent. V. Steinbach (Cologne University) drew attention to the potential role of upper-mantle phase transitions in governing the radial character of mantle convection in the large terrestrial planets⁹. He and D. Yuen argue that such phase transitions promote the formation of distinct convecting layers in the upper and lower mantle, but that as the Rayleigh number (governing vigour of flow) in the mantle decreases in response to core cooling, whole-mantle convection tends to become favoured over layered convection. If Venus cooled more rapidly than the Earth early in its history, it may now

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Jet Propulsion Laboratory

FIG. 1 Magellan radar image of an impact crater on Venus, approximately 70 km in diameter. Radar-bright, rough-textured ejecta extending up to 60 km from the crater rim and the remarkable bright flow features extending more than 300 km from the crater walls are thought to date from the impact¹ and have not been modified to any significant degree by later deformation or volcanism outside the crater (although the generally radar-dark, smooth floor of the crater may contain younger volcanic deposits).

*Fall meeting of the American Geophysical Union, San Francisco, 7–11 December 1992.

CELL MOTILITY

Variations on the theme of movement

Edwin W. Taylor

have a smaller mantle Rayleigh number and may have gone through a transition to whole-mantle convection, while the Earth may still be characterized by layered mantle flow. Such a transition would have been accompanied by overturn of the upper mantle, the upward transport of heat, and probably by global resurfacing.

The lithospheric instability models as applied to date for Venus suffer from being one-dimensional — global synchronicity is imposed by assumption. Even if such instability mechanisms do operate, it is likely that different parts of the planet will be at different stages in the stabilization and destabilization sequence, and whole-planet temporal variations will be smoothed out by such regional differences. For both lithospheric and mantle-convection mechanisms, attention needs to be given to all the terrestrial planets as a group. Both the Earth and Mars have gone through episodes of great magmatic activity, presumably an indication of variable mantle convective flux, but neither planet shows any evidence of rapid, complete global resurfacing during the past 4,000 million years, and nor do Mercury or the Moon.

Global resurfacing by large-scale overturn of the lithosphere or upper mantle should result in efficient outgassing of the upper mantle; consideration of the ^{40}Ar abundance of Venus's atmosphere suggests that any such widespread outgassing was restricted to times much earlier than 500 million years ago¹⁰.

Finally, remember that there are alternatives to the catastrophic model that seem to be equally consistent with the characteristics of impact craters on Venus². Further study of the Magellan images of impact features and continued development of dynamic models of the planet's interior should sharpen the competing hypotheses and may lead to an improved understanding of mantle convection and melting on all of the terrestrial planets, including the Earth. □

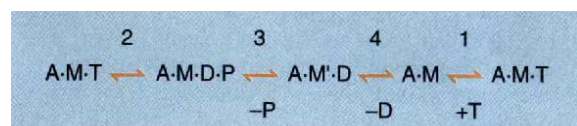
Sean C. Solomon is in the Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington DC 20015, USA.

THE discovery of the motor molecules kinesin and cytoplasmic dynein added two different motility systems to the well known actomyosin system. And with three members in a class one can begin to generalize. But although the idea that all actomyosin and microtubule motor systems are based on the same mechanism of chemomechanical coupling is an attractive one, there are puzzling differences between them. On page 168 of this issue¹, Romberg and Vale report studies of the effects of ADP and nucleotide analogues on kinesin-mediated motility that challenge the hypothesis that there is only a single kind of chemomechanical coupling.

The problem of how the force production in actomyosin contraction is linked to the enzyme hydrolysis cycles that generate the chemical energy for the process had proved intractable given that only one system was available for study. Once it was joined by the cilia dynein microtubule system, similarities in the kinetic scheme of actomyosin and the cilia system supported the idea of a unified theory². Kinesin ATPase, which is activated by microtubules, shares some kinetic properties characteristic of myosin, the force-generating protein in muscle which is an actin-activated ATPase: it has a phosphate burst and slow dissociation of one of the reaction products^{3,4}. It also has some similar structural features such as two heads and an α -helical coiled coil tail. But studies of microtubule motion driven by kinesin showed a puzzling difference from the actomyosin system, in that one motor molecule produced a somewhat larger velocity than many motors^{3,6}.

Romberg and Vale now find that ADP at a less than 1:1 ratio to ATP not only reduces the velocity of movement of microtubules driven by a single kinesin motor, but dissociates the microtubule from the surface. Low concentrations of ATP, below the K_m of the ATPase, or the slowly hydrolysed substrate, ATP- γ -S, prolonged the lifetime of the attached state more than tenfold while reducing the distance moved before detachment by only about twofold. Vanadate or aluminium fluoride, which may mimic an ADP-phosphate state, do not lead to detachment. All this suggests that kinesin-ATP and kinesin-ADP-P complexes are more strongly bound to microtubules than kinesin-ADP.

The results are surprising when compared with the 'classical interpretation' of actomyosin motility. In terms of the simplest kinetic scheme (see figure) force is developed by the transition to the A·M·D state which introduces a strain in a linear elastic element⁷. The strain or step size is taken to be 15 ± 5 nm. The force, in the positive or the negative direction, continues to be exerted as the actin filament moves past the myosin head until rapid detachment



Simple kinetic scheme for actomyosin. A, actin; M, myosin; T, ATP; D, ADP; and P, phosphate.

takes place with the formation of the A·M·T state.

This model accounts for the high velocity of actin movement ($5,000 \text{ nm s}^{-1}$ at 20°C) as the summation of the strain forces exerted by independent myosin heads each acting for a small fraction of the ATPase cycle time. Two important properties of the mechanism seem to be essential. First, the rate constant of ADP dissociation ($k_4 = 400 \text{ s}^{-1}$) is 20 times larger than the cycle rate⁸; thus the force can be exerted for 1/20th of the cycle time. Consequently, the velocity increases with the number of myosin heads interacting with the actin to a maximum value of k_4 times the step size⁹. Second, the model requires that myosin in the A·M·D state dissociates slowly and is more tightly bound than myosin in the A·M·T and A·M·D·P states, which must dissociate very rapidly (at least $5,000 \text{ s}^{-1}$) to minimize the drag exerted by these states.

The kinesin-microtubule system lacks the features that seem to be important for the actomyosin system. The strength of binding of the ATP and ADP complexes may be reversed and the rate constant of ADP dissociation is comparable to the cycle rate^{10,11}. The results of Romberg and Vale suggest that we may have to rethink the problem of how this system works. Has kinesin got the ATPase cycle coupled backwards compared to actomyosin? The kinesin motor moves towards the plus end of the microtubule while the *ncd* protein, which is closely related to kinesin, moves towards the minus end¹². If the coupling could be reversed, would the motor run in the opposite direction along the microtubule? Although the authors are cautious

- Schaber, G. G. *et al.* *J. geophys. Res.* **97**, 13257–13301 (1992).
- Phillips, R. J. *et al.* *J. geophys. Res.* **97**, 15923–15948 (1992).
- Shoemaker, E. M., Wolfe, R. F. & Shoemaker, C. S. *Lunar planet. Sci.* **22**, 1253–1254 (1991).
- Sandwell, D. T. & Schubert, G. *J. geophys. Res.* **97**, 16069–16083 (1992).
- Smrekar, S. E. & Phillips, R. J. *Earth planet. Sci. Lett.* **107**, 582–597 (1991).
- Parmentier, E. M. & Hess, P. C. *Geophys. Res. Lett.* **19**, 2015–2018 (1992).
- Stevenson, D. J. & Bittker, S. S. *Lunar planet. Sci.* **21**, 1200–1201 (1990).
- Arkani-Hamed, J. & Toksöz, M. N. *Phys. Earth planet. Inter.* **34**, 232–250 (1984).
- Steinbach, V. & Yuen, D. A. *Geophys. Res. Lett.* **19**, 2243–2246 (1992).
- Matsui, T. & Tajika, E. *Lunar planet. Sci.* **22**, 863 (1991).

in not giving a detailed interpretation of their results in the absence of kinetic data on this system, they consider the possibility that ATP binding may induce the transition that generates the force.

There are many questions to be answered before we can understand this system. The lack of dependence of velocity on motor number could be explained by the force generating state occupying a relatively large fraction of the cycle time and/or a large contribution to the drag from slow dissociation of the various cycle intermediates. Rate constants for dissociation of kinesin states from microtubules are not available and are crucial for the interpretation. The rate of dissociation and re-binding of single heads must be at least 50 to 100 s^{-1} in order to slip the kinesin along the microtubule, otherwise multiple motors would greatly reduce the velocity.

There is no simple explanation of the 11-s residence time obtained by Romberg and Vale, as this should correspond to 50 to 100 ATPase cycles, and kinesin with an ADP on both heads should come up much more frequently if the heads are independent. The distance moved per ATP is also a problem. In at least one case, ATPase and motility were measured in the same experiment and the distance per ATP per head was much larger than 10 nm^{13} . The problem of whether there is one-to-one coupling of the ATPase cycle to the mechanical cycle may be more serious for this system than for actomyosin. The properties of the kinesin-microtubule system may even fit the requirements of Yanagida's model¹⁴ better than the actomyosin system. □

Edwin W. Taylor is in the Department of Molecular Genetics and Cell Biology, The University of Chicago, Chicago, Illinois 60637, USA.

Transforming frogs and flies

Joan Marsh

IN *Metamorphosis*, published in 1916, Franz Kafka described the effects on a man of his transformation into an insect. More commonplace but no less intriguing phenomena, such as the metamorphosis of pupa into adult or tadpole into frog, were the subject of a recent meeting* of biologists. They gathered to debate the questions of which genes respond to the hormones concerned (ecdysteroids in invertebrates and thyroid hormone in vertebrates) and why some tissues regress while others grow or change. The underlying matter on most participants' minds, however, was whether there are parallels between the two hormonal regulatory systems.

Production of the hormones is the intermediate result of the triggering by environmental signals (light, temperature) of the release of neuropeptides, which in turn stimulate hormone release. But what happens after that? In *Drosophila*, ecdysone directly induces transcription of a small class of 'early' genes, including that for its receptor, and represses transcription of a large class of 'late' genes (M. Ashburner, University of Cambridge); this topic was covered in a recent News and Views article (*Nature* 357, 539; 1992).

A similar model can be proposed for amphibian metamorphosis (J. Tata, National Institute for Medical Research, London): thyroid hormone (TH) binds to a small amount of its receptor, probably derived from maternal messenger RNA in the oocyte, leading to activation of genes for its receptor and other 'early' genes with high-affinity sites for the TH receptor. Higher concentrations of receptor then induce transcription from 'late' genes encoding adult-specific proteins. Increased expression of the genes for the α - and β -TH receptors is indeed one of the earliest responses of the metamorphosing tadpole to the rise in endogenous TH (D. Brown, Carnegie Institution, Baltimore; B. Atkinson, University of Western Ontario; Y. Yaoita, Tokyo Metropolitan Institute for Neurosciences). Specific transcription factors may be involved in the earliest action of TH, one candidate being the CAAT enhancer-binding protein whose developmental expression in culture is closely correlated with that of the TH receptor genes.

Tissue-specific responses to ecdysone or TH range from cell proliferation to cell death, accompanied either by regression of the tissue or by its repopulation

with different cell types. The limb bud and the tail each consist of nerve, muscle and skin, but in *Xenopus* tadpoles TH stimulates growth of one and death of the other. Two 'immediate early genes' that are up-regulated in the tail during the first two hours after addition of TH, and are then down-regulated, have been identified (Brown). After 24 hours, some 30 mRNAs show increased levels in the tail and more than 100 in the limb bud. Fewer than 10 genes are estimated to be down-regulated in the tail in response to TH. In the limb, TH induces two waves of gene expression: first, of the direct response genes (some of which are also expressed in the tail), then of many genes associated with general cell growth and replication.

The transition of amphibia from an aquatic to a terrestrial existence is marked by dramatic changes in several organs. For example, larval erythroid cells die and are replaced by cells expressing the adult haemoglobin (R. Weber, University of Bern). This switch is stage specific and dependent on TH: if a mixture of larval and adult erythrocytes from a tadpole is treated with TH, the larval cells die and the adult ones survive (V. Galton, Dartmouth Medical School).

Changes also occur in the *Xenopus* immunological system at metamorphosis (L. Du Pasquier, Basel Institute of Immunology). In larvae, T cells are present but function poorly. MHC class I antigens are absent and the distribution of MHC class II is limited. Expression of class I antigens is turned on progressively throughout metamorphosis but in a way that may be simply age dependent, whereas the expression of class II antigens on all lymphocytes and the development of the adult B-cell antibody repertoire require metamorphosis.

Genes encoding liver-specific enzymes of the urea cycle are up-regulated by TH in the frog *Rana catesbeiana* (Atkinson). The interest here is that transcription of the genes is maintained after levels of TH and its receptor have fallen, illustrating a permanent switch to the adult expression pattern. Another gene to show this switch is that for the *Xenopus* adult-specific keratin (L. Miller, University of Illinois at Chicago); in cultured epidermal cells exposed transiently to TH before they become responsive, the gene 'remembers' the TH and is expressed five days later at the time normal siblings would turn on the gene.

Turning to insects, the cells destined to form the adult are segregated in discrete regions of the larva as imaginal

- Romberg, L. & Vale, R. D. *Nature* **361**, 168–170 (1993).
- Johnson, K. A. A. *Rev. Biophys. biophys. Chem.* **14**, 161–188 (1985).
- Hackney, D. D., Malik, A. & Wright, K. W. *J. biol. Chem.* **264**, 15943–15948 (1989).
- Sadhu, A. & Taylor, E. W. *J. biol. Chem.* **267**, 11352–11359 (1992).
- Howard, J., Hudspeth, A. J. & Vale, R. D. *Nature* **342**, 154–158 (1989).
- Block, S. M., Goldstein, L. S. B. & Schnapp, B. J. *Nature* **348**, 348–352 (1990).
- Hibberd, M. G. & Trentham, D. R. A. *Rev. Biophys. biophys. Chem.* **15**, 119–161 (1986).
- Taylor, E. W. *J. biol. Chem.* **266**, 294–302 (1991).
- Yueda, T. P. Q., Kron, S. J. & Spudich, J. A. *J. molec. Biol.* **214**, 699–710 (1990).
- Hackney, D. D. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6314–6318 (1988).
- Sadhu, A. & Taylor, E. W. *Biophys. J.* **59**, Abs. 567a (1991).
- Walker, R. A., Salmon, E. D. & Endow, S. A. *Nature* **347**, 780–782 (1990).
- Shimizu, T. et al. *J. Cell Biol.* **112**, 1187–1189 (1991).
- Harada, Y., Sakurada, K., Aoki, T., Thomas, D. D. & Yanagida, T. *J. molec. Biol.* **216**, 49–68 (1990).

**Metamorphosis*, Haslemere, Surrey, 4–7 October 1992.

disks — sacs of embryonic epithelium. An imaginal leg disk cultured in 20-hydroxyecdysone produces an adult leg in 18 hours (J. Fristrom, University of California at Berkeley): this does not involve cell proliferation, merely elongation of individual cells driven by ecdysone-dependent contraction of an actin-myosin ring. The actin and myosin genes are not directly affected by ecdysone; the hormonal effects are mediated by genes of the Broad-Complex.

The action of ecdysone and TH are countered by juvenile hormone (JH) and prolactin respectively, the balance of the two hormonal systems being important in determining the timing of the onset of metamorphosis. For instance, JH regulates the synthesis of the cuticle proteins in the moth *Manduca sexta* (L. Riddiford, University of Washington, Seattle). It has no effect on the ecdysteroid-induced expression of genes required for new cuticle at moults associated with growth but prevents activation of genes specific for the transition to pupa or adult. The effects of 20-hydroxyecdysone on the cuticular protein genes are indirect. A possible mediator of the ecdysteroid induction is MHR3, an orphan steroid receptor, which may regulate stage-specific cuticular protein genes but has no effect on at least one of the genes for enzymes required for cuticular sclerotization after ecdysis. A putative JH receptor has been identified, but JH action remains a grey area.

In *Xenopus*, prolactin abolishes transcription of both early and late genes, those for the TH receptor and for keratin and albumin. It also has growth-promoting and osmoregulatory effects, implantation of pellets impregnated with prolactin in the larval spleen (from which blood drains to the liver) resulting in marked tail growth in *R. catesbeiana* (C. Nicholl, University of California at Berkeley). The hepatic factor that synergizes with prolactin, named synlactin, has not yet been purified. Insulin-like growth factor (IGF-I) can mimic synlactin in one biological assay, but it has not been found in frog serum (although an IGF-binding protein has).

So where does that leave us? The genes that respond to hormonal signals are being identified and the tissue-specific responses can be explained in terms of expression of hormone receptor isoforms and the requisite transcription factors. But the differences between the processes of metamorphosis in insects and vertebrates suggest that similar events are not mediated by homologous gene products. Kafka might not have been surprised. □

Joan Marsh is at the Ciba Foundation, 41 Portland Place, London W1N 4BN, UK.

Sorting out natural stone stripes



MUNDANE geomorphic processes have an uncanny ability to create remarkable patterns in the world around us. Bed-forms such as ripples or dunes formed under flowing air or water, beach cusps, evenly spaced rills on hillsides, and meanders in rivers are just a few examples. On page 142 of this issue, B. T. Werner and B. Hallet use computer cellular automata to model the development of sorted stripes of stones, shown above from near the summit of Mauna Kea, Hawaii. Along with sorted circles, these are among the most striking patterns in arctic and alpine landscapes. In none of these cases is there an external template to impose the pattern, which instead must be entirely self-organized, bringing investigation of them into an arena of more general interest in physics and chemistry.

In the past, perturbation methods have been applied to such studies to find the fastest-growing perturbation. But the fastest-growing initial perturbation is not always equivalent to the final wavelength, because the pattern becomes strong enough for local finite-amplitude effects to dominate the physics. As many interesting geomorphic problems involve the rearrangement of individual grains and are the result of independent motions modulated only by local conditions, others have turned to cellular-automaton techniques. In these, the complex physics is boiled down to a simple set of rules that is applied to every grain in the system. Individual grains, represented by pixels on the screen and perhaps coloured to represent different grain sizes, are moved from one position to another according to these rules. The power of the technique is that subtly different rules may be easily tested to reveal the relative importance of one or another feedback in creating the patterns.

The sorted stripes that Werner and Hallet investigate form on unvegetated arctic and alpine hill slopes subject to

periodic freezing. Under these circumstances, stones larger than several millimetres become separated from the frost-susceptible fine-grained domains to create a pattern of stone stripes aligned up- and downslope. On lower slopes, the same physics apparently creates sorted nets. Hallet brings to bear considerable experience in both theory and observation of periglacial phenomena, while Werner has helped to pioneer the application of cellular-automata and grain-dynamics techniques to geomorphology. The process they finger as the cause of sorted stripes is ice needling, wherein tiny pillars of ice, sometimes several centimetres long, form beneath the surface grains on cold, clear nights. When the needles thaw, the grains are lowered back to new positions on the surface, typically downslope.

Several feedbacks operate in the growth of the stone stripes. Most important is the differential frost heave in fine- and coarse-grained domains. Ice needling transports the coarser grains — the stones — on average perpendicular to the local contour, expelling them from the fine-grained domains. But the coarser domains are less susceptible to heaving, so that ice needling is less likely to drive the stones back. Subtle perturbations in an otherwise random walk therefore lead to this strikingly regular textural pattern. In the absence of a mean slope, the simulations produce sorted nets, passing yet another test against the natural landscape. The authors not only match the statistics of the stripe spacing, they also point to new measurements that may pinpoint the physics more tightly; downslope motion of grains should be faster within the fine domains, for example.

Robert S. Anderson

Robert S. Anderson is in the Earth Sciences Department, University of California, Santa Cruz, California 95064, USA.

Sex to order

A BIRTH announcement with a difference is made by D. G. Cran and colleagues in the latest issue of *The Veterinary Record* (**132**, 40–41; 1993) — the production of three female and three male calves from eggs fertilized by spermatozoa preselected for sex. Behind the news, which in the long run will be good news for beef and dairy farmers, is the refinement of a flow cytometry technique for sorting sperm according to whether they contain X or Y chromosomes. Use of a fluorescent dye and laser detection means that the 4 per cent or so difference in DNA content between X- and Y-bearing sperm can be distinguished with reasonably high efficiency. The birth (after embryo transfer into heifers) of the calves with sexes as predicted was the test that the sperm remain viable after their buffeting, but it will be some time before sexed embryos become available commercially. The necessary large-scale trials have yet to be completed.

Exclusion principle

TRACKING the motion of iridium atoms over the surface of an iridium crystal, S. Wang and G. Ehrlich find a curious *cordon sanitaire* just three atoms wide around a surface defect (*Phys. Rev. Lett.* **70**, 41–44; 1993). The defect, deliberately introduced, is a small cluster of iridium atoms which mimics the step that forms on a surface as a crystal grows. Alternately freezing the crystal to locate the atoms with field-emission microscopy and then warming it to allow them to wander, the authors find that although the atoms might move anywhere just beyond the exclusion zone, they are never seen inside it. Eventually, the lone atoms can suddenly join the central cluster. What keeps them out of the exclusion zone is not clear — tests seem to rule out the presence of an energetic barrier.

Myosin multiplicity

Acanthamoeba contains three forms of myosin I, the tailless monomeric myosin that forms no filaments. I. C. Baines *et al.* (*J. Cell Biol.* **119**, 1193–1203; 1992) have now determined their distribution in the cell and inferred something of their functions. From its location it seems that myosin IA acts in nipping off the small plasma membrane vesicles and moving them between the membrane and cytoplasm; IB turns up in the fraction of large vacuoles and phagocytic vesicles that derive from the active membrane regions, and thus most probably drives the extension of the membrane; and IC is the only form seen in the contractile vacuole and so is probably involved in its function. It will be especially interesting to discover how these myosins recognize their chosen sites.

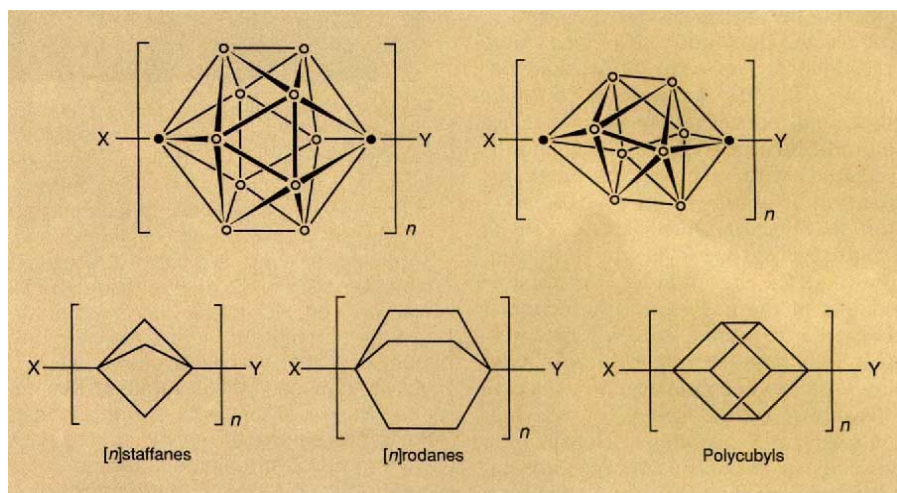
Carborod molecular scaffolding

Jeffrey S. Moore

As miniaturization of electronic, photonic and mechanical devices approaches the molecular scale, chemists are recognizing the possibility of building the smallest possible objects from the atom up. This creates the need for the molecular construction kit, which would probably include an assortment of molecular sticks of different lengths, connectors, and a mechanism to bring the pieces together. But even the simplest component of this kit, the non-collapsible rod,

carborane cage means that the carborods uniformly fill space around their longitudinal axis. Thus, of all the so-called 'rigid-rod' molecules, the carborods most closely resemble cylindrical pillars. They are available in lengths that are integral multiples of about 4.5 Å, the average repeat distance of the carborane monomer.

Besides their geometry, carborane monomers have other characteristics that make them ideal constituents



Molecular scaffolding: top, the two kinds of carborod just synthesized; below, three rigid rods previously synthesized. Solid circles, carbon; open circles, boron (B—H).

has proved pretty elusive. To help fill this gap, independent groups in the laboratories of M. Frederick Hawthorne and Josef Michl have developed a new family of stiff molecular scaffolding, named carborods (X. Yang *et al.*, *J. Müller et al. J. Am. chem. Soc.* **114**, 9719–9721, 9721–9722; 1992).

These new rods are oligomers of *para*-carborane monomers. Carboranes are inorganic cage molecules with boron and carbon vertices. In the *para* isomers, the carbon atoms occupy the apical positions of high-symmetry polyhedra (see figure). Both laboratories report carborods using a 12-vertex carborane monomer having icosahedral symmetry. Michl's group also described a carborod based on a 10-vertex carborane monomer that has a bicapped square antiprism geometry. The carbon atom sites in both of these polyhedra are ideally situated for linear extension through polymerization.

Carborods are thick compared to known rigid-rod molecules (such as so-called $[n]$ staffanes, $[n]$ rodanes and polycubyls, see figure). Their van der Waals diameter is estimated to be more than 7 Å. Moreover, the high symmetry of the

for molecular scaffolding. They are exceedingly robust, having exceptional thermal and chemical stability. They are transparent to ultraviolet and visible radiation, a feature that will be most advantageous in some applications. And perhaps most importantly, carborane chemistry provides a means to prepare rods of discrete length and with specified end-group functionality. This is possible because of the acidity of the terminal CH groups, which offers site-selective reactivity in the growing, multi-vertex carborods.

A practical molecular construction kit will require components with precisely defined structure. Along with uniform length, controlled end-group functionality will be needed to link the rod with connector pieces, through covalent or non-covalent bonding. In making rods of uniform length, one must be able to control the oligomerization reactions. Many of the molecular rods that have been previously described do not allow a means to do this. The methods reported by Hawthorne's and Michl's groups allow carborods to be obtained with modest control. They succeeded in preparing and characterizing end-

RIBOZYMES

functionalized rods uniquely possessing one, two, three and four carborane repeat units. A single-crystal X-ray structure from Hawthorne's group shows the tetramer to be 17.21 Å in length and almost perfectly linear. This is apparently the longest rigid-rod molecule to be structurally characterized.

The components of a practical molecular construction kit will need to be soluble in common solvents. This might seem like a trivial issue, but chemists as molecular constructors perform their building operations in solution, and rigid rods are notoriously insoluble. Problems of solubility were apparent with the new carborods as well. Nonetheless, Hawthorne's group shows that the poor solubility of the parent tetramer can be overcome through proper choice of end-groups. They hope to use these tactics in combination with an efficient, repetitive synthetic scheme to construct even larger rods.

It is popular to draw analogies between a child's tinkertoys and molecular construction kits, but this overlooks one important difference. Whereas the child can physically handle the components to construct an object piece by piece, to attempt the same in building a molecular object of even moderate size would be cumbersome. The components of a practical molecular construction kit will have to be smart enough to assemble themselves automatically into the desired arrangement. Herein lies the greatest challenge — how to encode the pieces of scaffolding appropriately. This might involve adding extra groups to the pieces, to bias their orientation and position. The chemistry of the non-covalent bond will surely play an important part in this.

Michl's group's effort to produce molecular tinkertoys has already begun to confront the challenge of automatic assembly. Their study included the preparation of Langmuir-Blodgett films — monolayers on a water surface — using carborod building blocks. The carborods were encoded with a carboxylic acid group at one end, a constraint intended to define the orientation of the carborod at the water-air interface. The other end contained a reactive iodide group. The area per molecule of the films was consistent with a monolayer in which the rods packed end-on at the interface. If all goes according to plan, this should leave a layer of exposed iodine atoms. The next step will be to activate the iodine atoms photochemically to allow continued molecular construction on this surface. Chemists will then truly become molecular builders. □

Jeffrey S. Moore is in the Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109, USA.

Evolution *ex vivo*

Jack W. Szostak

BIOLOGICAL evolution is a slow process. But populations of replicating macromolecules can undergo evolution *in vitro* over a timescale more suited to human observers. On page 182 of this issue¹, Lehman and Joyce apply techniques of directed *in vitro* evolution² to modify the catalytic requirements of that much-studied ribozyme — the *Tetrahymena* ribozyme — with some striking results.

The challenge is daunting; how can we obtain ribozymes with specified properties in the absence of knowledge and understanding of RNA structure sufficient to permit rational design? The answer, in principle, is simple: make all possible changes in a given ribozyme (or at any rate as many changes as are practical, which turns out to be a great many), and then select for the particular changes that alter the properties of the ribozyme in the desired manner. Darwinian evolution — the preferential survival and reproduction of the fittest variants in a population — applies just as much to these replicating populations of informational macromolecules *in vitro* as it does to living organisms on a large scale over long periods of time.

This approach was first applied by Beaudry and Joyce³ to the evolution of a variant form of the *Tetrahymena* ribozyme with altered substrate specificity. Starting with the wild-type *Tetrahymena* ribozyme, which prefers RNA as a substrate, they were able to select for a variant ribozyme with a 100-fold increase in activity on a DNA substrate. This was accomplished by developing methods of mutagenesis, amplification, and especially selection, for the molecules with the desired properties (Fig. 1). The cycle of mutagenesis, selection and amplification was repeated for 10 'generations', at the end of which individual molecules were analysed to determine the mutations responsible for the increase in activity.

In the new work of Lehman and Joyce¹, the same methods have been applied to a new task — changing the metal ion specificity of the ribozyme. The ribozyme requires divalent cations for two distinct roles, folding and catalysis. Although either Ca^{2+} or Mg^{2+} will work at most sites, there is at least one, probably the active site, at which Ca^{2+} will not work⁴. Piccirilli *et al.* have recently shown⁵ that a Mg^{2+} ion interacts directly with the 3' oxygen of the phosphodiester bond that is attacked during the catalytic reaction. This Mg^{2+}

is presumed to stabilize the developing negative charge on the 3' oxygen during chain cleavage. The implication is that an important part of the mechanism of ribozyme catalysis involves binding this Mg^{2+} and directing it towards the appropriate 3' oxygen. The ribozyme- Mg^{2+} interactions involved are not known.

Lehman and Joyce started with a large pool of mutant ribozymes, 2.4×10^{13} by their reckoning, generated by a 5 per

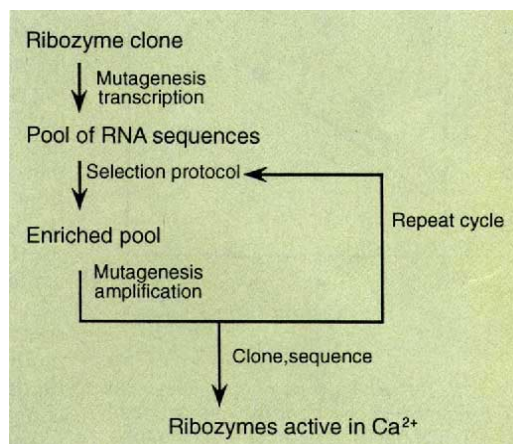


FIG. 1 Cycle of ribozyme mutagenesis, selection and amplification.

cent mutagenesis over 140 nucleotides of the catalytic core of the ribozyme. Such a pool is sufficient to include all five-base changes in the enzyme sequence, and about half of the possible six-base changes. They then selected for mutant ribozymes that could function in the presence of Ca^{2+} with no added Mg^{2+} , using the same selection procedure as in their previous experiment, so that only catalytically active molecules could be replicated. Competition in the Joycean jungle is severe — no more than 0.1–1 per cent of the population survives the selection phase and is allowed to reproduce to form the next generation. Each survivor, however, generates 100–1,000 progeny. Additional mutations were introduced during the amplification phase of each generation by the use of a mutagenic polymerase chain reaction.

One of the exciting aspects of *in vitro* evolution is the ability to assess the changing nature of the population through time. Lehman and Joyce sequenced 50 clones from generations 2, 4, 6 and 8, to see the results of their selection protocol. The molecules in the initial pool contain an average of seven mutations each, as a consequence of the initial 5 per cent mutagenesis over 140 nucleotides. After two generations of

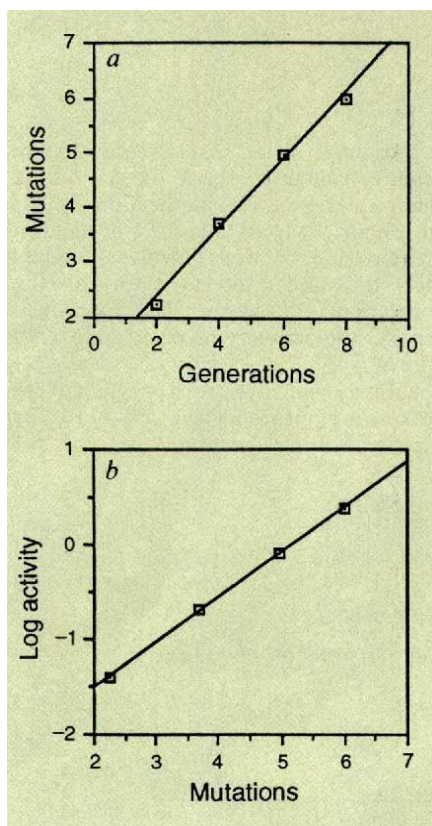


FIG. 2 *a*, The average number of mutations per sequence in the pool increases linearly with the number of generations. *b*, The log of the activity of the pool is a linear function of the average number of mutations per sequence in the pool.

selection, the average number of mutations in the surviving sequences is only 2.26; clearly, the vast majority of mutations are harmful and have been eliminated. Of course, clones with many more mutations are still present — they are just too rare to be seen. After this initial phase, the progress of the selection is remarkably uniform, with exponentially increasing activity and a steady linear increase in the average number of mutations (Fig. 2*a*). The gradual increase in the average number of mutations is due largely to the enrichment of progressively more active, but initially rarer, sequences that were present in the initial population. If a superior mutant sequence were present in the initial population, it would increase rapidly in numbers, and the subsequent creation of the same mutant sequence by new mutations would not contribute significantly to its abundance. Although the effect of new mutations is probably small after only eight generations, it seems quite likely that continued selection will lead to variants with so many beneficial mutations that they would not have been present in the original population.

The linear relationship between the log of the pool activity and the average number of mutations per sequence is

surprising (Fig. 2*b*). The simplest interpretation of this correlation is that all the mutations contribute equally and independently to the increasing activity; why this should be so is less clear. This plot also tells us something about the shape of the activity 'hill' in sequence space: there seems to be at least one, and perhaps several, easily climbed ridges that lead gradually to a peak of activity. The fact that two of the mutations are found in mutually exclusive subsets of clones suggests that there are at least two different solutions to the problem of functioning with Ca^{2+} . The location and height of these peaks must await further generations of exploration.

How do the selected mutations change the ribozyme so that it can work with Ca^{2+} alone? It is by no means obvious how one could change a metal-ion binding site by simply changing the bases involved in direct interactions. Lehman and Joyce argue strongly from kinetic data and measured metal ion concentrations that the enzyme has not simply become so adept at binding Mg^{2+} that it can function in the low but unavoidable contaminating Mg^{2+} . Two more interesting possibilities are that the Mg^{2+} -specific site of the ribozyme has become modified so that it can bind either Mg^{2+} or the larger Ca^{2+} ion (radius 100 pm, compared with 72 pm for Mg^{2+}), or that it has become able to function without any divalent cation at all.

It is unlikely that any of the selected mutations are part of the critical metal-binding site of the ribozyme. Inspection of the three-dimensional model of Michel and Westhof⁶, which is consistent with a remarkable number of other experimental tests, reveals that none of these mutations is close to the reactive centre. It seems more likely that they are at distant locations, and exert their effects by causing subtle changes in the geometry of the ribozyme, for example changing the twist of a helix, or the angle between two helices, the effects of which are then propagated to the binding site to displace the directly interacting bases to positions in space consistent with the binding of both Mg^{2+} and Ca^{2+} . Obviously, these experiments are crying out for a high resolution X-ray crystal structure for a proper interpretation. □

Jack W. Szostak is in the Department of Molecular Biology, Massachusetts General Hospital, and the Department of Genetics, Harvard Medical School, Boston, Massachusetts 02114, USA.

1. Lehman, N. & Joyce, G. F. *Science* **261**, 182–185 (1993).
2. Szostak, J. W. *Trends Biochem. Sci.* **17**, 89–93 (1992).
3. Beaudry, A. A. & Joyce, G. F. *Science* **257**, 635 (1992).
4. Celander, D. W. & Cech, T. R. *Science* **251**, 401 (1991).
5. Piccirilli, J. A., Vyle, J. S., Caruthers, M. H. & Cech, T. R. *Nature* **361**, 85–88 (1993).
6. Michel, F. & Westhof, E. *J. molec. Biol.* **216**, 585–610 (1990).

DAEDALUS

Second thoughts

LAST week Daedalus proposed his encephalomagnet, a cunning quadrupolar affair with hollow pole pieces. Placed over the subject's head, it induces an intense local ring current in a small region of his brain. He then forcefully recalls whatever memory is stored at that location.

In effect, this is a sort of local electroconvulsive therapy (ECT). Instead of passing through much of the brain, with widespread disruptive effects, the current is neatly localized. Despite its fearful reputation, ECT relieves depression very effectively, though it does seem to cause some loss of memory. Encephalomagnetic therapy should be much more controllable.

Why does ECT cause memory loss? Memory is said to be stored in the synapses where one neuron connects with another. The triggering neuron releases a neurotransmitter into the synapse; this has some finite probability of launching a nerve impulse up the dendrite leading to the receiving neuron, which fires in its turn. The synapse 'learns' by altering this probability, so that all our knowledge is ultimately stored as a huge matrix of probabilities. Now ECT fires neurons by direct electrical stimulation. The resulting nerve impulse must travel out to all the ramifications of the neuron — even down the dendrite back to the receptor synapse, an event that never happens naturally. This unnatural backwards impulse, says Daedalus, must erase the firing probability stored at that synapse. Hence the loss of memory.

Memory loss is not altogether a bad thing. Daedalus recalls that dreadful verdict on the Bourbons: "They learnt nothing. They forgot nothing". Many creative people lose their spark in middle age: they know all the answers, but they don't notice that the questions have changed. So Daedalus's encephalomagnetic therapy may be a powerful rejuvenator. The local ring current could be used to explore round the brain until it stimulates some obsession, rigid misapprehension or area of total complacency. A cautious increase of current will then start to reset the local neurons, erasing their firing probabilities and freeing them to learn anew. The subject will find his certainties softening; new questions and ideas will begin to encroach on his previous unalterable convictions. He could even control the whole process himself. Artists, scientists, engineers and creative people generally should welcome this opportunity to renew their mental landscape. Sadly, the politicians, demagogues, academics and company chairmen most in need of the treatment will resist it fiercely.

David Jones

Rhino conservation tactics

SIR — Although African elephant populations are increasing as a result of the moratorium on the ivory trade begun in 1989, the signs are less encouraging for black rhinos, whose prodigious decline continues¹. Efforts aimed at the conservation of Africa's black and white rhinos have increasingly taken three forms — translocation to safe reserves, attempts to halt the illicit market, and dehorning, a controversial procedure first tried in Namibia in 1989 (refs 2,3). Dehorning is still practiced there and it has been adopted in Zimbabwe on both species. The efficacy of this programme remains uncertain, although data gathered by my colleagues and me from a recent field study of black rhinos in Namibia's Kunene River Province and Etosha National Park offer some insights.

First, regeneration of rhino horn is rapid, averaging nearly 9 cm of total horn per animal per year (Fig. 1). Because the estimated black-market value of newly grown horns ranges from \$1,775 to \$7,750 on an average animal just 1 year after dehorning⁴ and removal costs about \$1,400 per animal, the price of annual horn pruning is exceeded by the horn's black-market value in just 10 months. It was also concluded in a recent modelling exercise that intensive horn harvesting must be adopted if rhinos are to remain valueless to poachers⁵.

Second, this finding is reinforced by new information on the choices of

poachers. If larger-horned rhinos were preferred trophies, then recently dehorned individuals would remain of low value. However, the similarity between mass categories of horns confiscated from poachers and those on live rhinos in extant, wild populations (Fig. 2)



FIG. 1 Regrowth of horns in an adult black rhino female 3 years after dehorning.

points to a lack of discrimination by poachers. This suggests that even rhinos with small horn nubs are vulnerable to poachers.

Third, although the claim made regularly by the popular media, that dehorning must be working since no hornless rhinos have been poached, seems encouraging, it is also true that horned Namibian black rhinos have not been

poached either. In reality, it is impossible to conclude anything about the demographic effects of dehorning. From a behavioural perspective, maternal responsiveness to dangerous predators such as lions and spotted hyenas has less to do with horn size than to the ages of the young. Mothers of calves less than 18 months old flee about five times more often than when the calves are older (J. B. and C. Cunningham, unpublished data).

Although these results might dampen enthusiasm for dehorning, any definite conclusions would be premature. One unforeseen consequence of dehorning has been the massive and often sensational media coverage. Poachers may have simply moved elsewhere or be waiting for the interest to die down. And political actions such as Namibian independence, the removal of South African defence forces, or the cessation of civil strife involving Angola, Namibia and South Africa, may have contributed to a demise in poaching. Whatever the cause(s), involving local people in the search for realistic ways to maintain viable populations of Africa's rhinos remains the critical challenge.

Joel Berger

*Program of Ecology, Evolution and Conservation Biology,
University of Nevada,
Reno,
Nevada 89512, USA*

1. Western, D. *Trends ecol. Evol.* **2**, 343–346 (1987).
2. Leader-Williams, N. *Nature* **340**, 599–600 (1989).
3. Lindeke, M. S. *Afr. J. Sci.* **86**, 226–227 (1990).
4. Milliken, T. et al. *Traffic Bull.* **12**, 17–21 (1991).
5. Milner-Gulland, E. J. et al. *Proc. R. Soc. Lond.* **B249**, 83–87 (1992).

A word in your protein

SIR — What is the longest word spelled out in the sequence of a protein in the protein sequence database¹ using the one-letter code for amino acids? None of the extensive literature devoted to this problem^{2–4} has taken a truly systematic approach; the longest word found to date contains only seven letters. We have matched the entire *Oxford Unabridged English Dictionary* (second edition, 20 volumes, 572,728,830 characters, with information content close to that of the human genome) against the entire SwissProt protein sequence database (version 23)⁵. Using the Patricia tree data structure¹, the matching consumed only 23 minutes of computational time. We found two words with nine characters: “hidalgism” (the manner or practice of a hidalgo) entered the English dictionary via a citation from 1887, and appears at positions 247–355 of the integrase of bacteriophage lambda (acquisition number P03700). “Enslists” (the plural of enslist, one who preserves

his crops by ensilage) entered the dictionary via a citation from 1883, and appears at positions 81–89 of the PRRB protein from *Escherichia coli* (acquisition number P17222).

In addition to being the longest strings appearing simultaneously in the English and protein languages, these are also candidates for the most unusable pieces of information simultaneously in lexicography and in biochemistry. Their discovery does, however, demonstrate the power of these data structures in handling large amounts of information.

Gaston H. Gonnet

Steven A. Benner

*Institute for Scientific Computation, and
Institute for Organic Chemistry,
ETH, Zurich, CH-8092 Switzerland*

1. Gonnet, G. H., Cohen, M. A. & Benner, S. A. *Science* **256**, 1443–1445 (1992).
2. Price, N. C. *Trends biochem. Sci.* **12**, 349 (1987).
3. Purton, M. *Trends biochem. Sci.* **13**, 48 (1988).
4. Jimenez, A. *Trends biochem. Sci.* **14**, 14 (1989).
5. Bairoch, A. & Boeckmann, B. *Nucleic Acids Res.* **20**, 2019–2022 (1992).

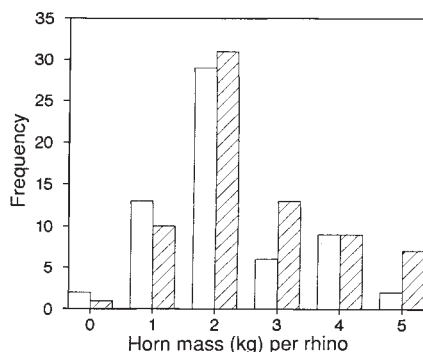


FIG. 2 Frequency distributions of the mass of horns (anterior and posterior combined) on adult and subadult black rhinos confiscated from poachers (open boxes) or on live rhinos (hatched boxes). Horn mass (Y) of live rhinos was estimated photogrammetrically based on measurements from 104 confiscated horns, where $Y = 15.49X - 0.21$ and X is horn circumference times length ($r^2 = 0.83$; $P < 0.0001$). Distributions of the poached and live samples do not differ (Kolmogorov–Smirnov test; $D = 0.444$).

Neurotoxicity of β -amyloid

SIR — The neuropathology of Alzheimer's disease is characterized, in part, by the deposition of abnormal protein aggregates called β -amyloid. The metabolism of the β -amyloid precursor protein has emerged as a central issue in pathogenesis because of the discovery of point mutations in its structural gene in some forms of early-onset familial Alzheimer's disease¹.

Yoshikawa *et al.*² extended our observations of β -amyloid precursor protein (APP)-mediated neurotoxicity, using the identical cell culture system (P19 embryonic carcinoma)³. Although neurotoxicity of APP derivatives therefore seems to be well established, there are substantial differences between the results of the three laboratories that have reported on the overexpression of complementary DNA (cDNA) constructs for APP, or for parts of APP, in cultivated neuronal cells²⁻⁴. These various observations should also be put in the context of more recent results.

(1) Which products of APP metabolism are toxic to cultured neurons? Neurotoxicity of a carboxy-terminal fragment (spanning the β -amyloid peptide and cytoplasmic domains of APP) was originally reported by Yankner *et al.*⁴ by overexpression of its cDNA in differentiated neurons derived from a PC12 transformant. These authors did not observe neurotoxicity of protein products from a full-length cDNA construct for APP-695, however, in contrast to Yoshikawa *et al.*². Using an antibody raised against the first 15 amino acids of the β -amyloid peptide, Yankner *et al.* observed high-molecular-mass (>110,000) protein aggregates on immunoblots from conditioned medium of a PC12 transformant bearing the DNA construct for the carboxy-terminal fragment. Fukuchi *et al.*³ and Yoshikawa *et al.*² did not observe such high-molecular-mass aggregates using antibodies against a cytoplasmic domain of APP in P19-derived neurons. Fukuchi *et al.*³ and Yoshikawa *et al.*² found multiple carboxy-terminal fragments, and Fukuchi *et al.* found that neuronal degeneration was associated with the generation of *M*_r 16,000 and 14,000 carboxy-terminal fragments.

(2) Is the neuronal specificity of the APP-derived cytotoxic moiety related to variations in processing or to variations in target cell specificity? Yoshikawa *et al.* did not find multiple carboxy-terminal proteolytic fragments of APP in muscle cells derived from P19 transformants, suggesting that its proteolysis is different from that in neurons. We have also induced differentiation of our P19

transformants to muscle cells by treatment with dimethyl sulphoxide, and find that the proteolysis of APP in the muscle cells is similar to that in the neuronal cells, but that the 16,000 and 14,000 products did not affect the survival of muscle cells. This discrepancy may be due to differences in the choice of promoters with resultant differences in levels of transgene expression. Using several different cell lines⁵, promoter activities of a cytomegalovirus enhancer and chick β -actin promoter (which was used for the construction of our expression vector) are much higher than those of a SV40 promoter, which was used by Yoshikawa *et al.* Furthermore, when a DNA construct for the carboxy-terminal fragment was introduced into a morphologically heterogeneous neuroblastoma cell line (SK-N-SH) (consisting of neural and non-neural cell types), selective neurotoxicity of the fragment was observed⁶.

(3) Could neurotoxicity be mediated by a secreted β -amyloid peptide? Recent observations^{7,8} have raised the possibility that neurotoxicity could be mediated by secreted, soluble β -amyloid peptide. To address this question, we have repeated the experiments of Fukuchi *et al.*^{3,6}, using conditioned media from stable transformants of P19 and SK-N-SH. Addition of conditioned media did not influence the neuronal differentiation of nontransfected P19 and SK-N-SH cells nor the survival of P19-derived and SK-N-SH-derived neurons. We also did not observe secreted 16,000 and 14,000 fragments in conditioned media from the P19 transformants. We therefore propose that intracellular accumulations of the 16,000 and 14,000 fragments result in neural degeneration.

Ken-ichiro Fukuchi

Bryce Sopher

George M. Martin

*Department of Pathology,
University of Washington,
Seattle, Washington 98195, USA*

YOSHIKAWA REPLIES — Yankner *et al.*⁴ and Fukuchi *et al.*³ found degenerative changes of neurons overexpressing a carboxy-terminal fragment of APP, which spans the β /A4 and cytoplasmic domains (the fragment lacks a signal sequence for secretion and membrane integration). Two years ago we demonstrated that a transient overexpression of the carboxy-terminal fragment, designated APP-C100, leads to a dense accumulation of amyloid-like fibrils near the nuclear membrane of COS cells⁹. Since then we have attempted to transfect APP-C100 cDNA into various cultured cell lines including PC12, P19, glioma and fibroblasts to establish stable transformants, but we failed to obtain any clones expressing high levels of APP-C100. Therefore, we assumed that an

intracellular overproduction of APP-C100 is extremely toxic to any type of cells (I suggest that the stable transfectants obtained in other laboratories^{3,4} express non-toxic levels of carboxy-terminal fragments under undifferentiated conditions). We transfected human full-length APP cDNA into P19 embryonal carcinoma cells, which differentiate into post-mitotic neurons following retinoic acid treatment, to examine whether overexpressed APP undergoes aberrant processing to generate APP-C100 or similar amyloidogenic fragments in differentiated neurons. We found severe degenerative changes of P19-derived neurons which contained high amounts of potentially amyloidogenic carboxy-terminal fragments².

More recently, we established human glioma cell lines (Bu-17) overexpressing full-length APP cDNA¹⁰. These transfectants accumulate carboxy-terminal fragments of APP and show degenerative changes when treated with chloroquine, a compound reducing lysosomal functions. Therefore, carboxy-terminal fragments of APP generated in the lysosomes may be toxic to non-neuronal cells as well as to neurons. We speculate that the susceptibility to overexpressed APP varies from cell to cell (even among neuronal subtypes) owing, at least in part, to the capability of metabolizing APP into non-toxic peptides. This may explain the absence of neurotoxicity of PC12 cells transfected with full-length APP cDNA⁴.

The carboxy-terminal fragments generated in the full-length APP cDNA transfectants are localized in the membrane fractions². Neither conditioned media of these APP cDNA transfectants (P19, Bu-17) nor APP-C100 purified from cDNA-transfected cells showed neurotoxicity when added extracellularly (Y. Hayashi & K. Y., unpublished observations). Therefore, these transfectants may be vulnerable to the intracellular amyloidogenic carboxy-terminal fragments.

The most likely idea is that extracellular amyloid deposits induce neuronal degeneration 'from outside'. Our cell models^{2,9,10} suggested a new concept, that intracellular amyloidogenic fragments kill neurons 'from inside'. In Alzheimer's disease and related disorders such as Down's syndrome, the

- Goate, A. *et al.* *Nature* **349**, 704–706 (1991).
- Yoshikawa, K., Aizawa, T. & Hayashi, Y. *Nature* **359**, 64–67 (1992).
- Fukuchi, K. *et al.* *Biochem. biophys. Res. Commun.* **182**, 165–173 (1992).
- Yankner, B. A. *et al.* *Science* **245**, 417–420 (1989).
- Miyazaki, J. *et al.* *Gene* **79**, 269–277 (1989).
- Fukuchi, K. *et al.* *Molec. Brain Res.* **16**, 37–46 (1992).
- Seubert, P. *et al.* *Nature* **359**, 325–327 (1992).
- Haass, C. *et al.* *Nature* **359**, 322–325 (1992).
- Maruyama, K., Terakado, K., Usami, M. & Yoshikawa, K. *Nature* **347**, 566–569 (1990).
- Hayashi, Y., Kashiwagi, K. & Yoshikawa, K. *Biochem. biophys. Res. Commun.* **187**, 1249–1255 (1992).

overproduction of APP and/or reduction of lysosomal activity might accumulate the cytotoxic fragments within neurons, which eventually degenerate and die. In my view, 'extracellular' amyloid plaques could be harmless traces of dead neurons in the brain of Alzheimer's disease and related disorders.

Kazuaki Yoshikawa

*Department of Molecular Neurobiology,
Tokyo Metropolitan Institute for
Neuroscience,
2-6 Musashidai,
Fuchu, Tokyo 183, Japan*

Cooling in the late Cenozoic

SIR — Raymo and Ruddiman in their Review Article¹ propose an explanation for the late Cenozoic climate cooling. Their approach, built from a comparison of data trends, is an inspired beginning. But it remains ultimately conceptual and primarily verbal, and thus omits an essential test of assembling a matter-conserving, dynamic numerical system. Because Raymo and Ruddiman contrast their concept (global weathering rate depends on rock supply) with opposing results drawn from a variety of quantitative models (global weathering depends on CO₂ supply), it is crucial that their hypothesis be subjected to similarly quantitative tests. Quantitative models of the carbonate-silicate geochemical cycle require an explicit system of equations for the fluxes of carbon to and from the ocean-atmosphere system and for atmospheric CO₂. I believe that if the authors were to place their hypothesis in the framework of a quantitative system model, major difficulties would appear.

Raymo and Ruddiman claim that uplift of the Tibetan plateau over the past 40 million years (Myr) provided a source of highly-weatherable silicate rock, which enhanced the rate of removal of atmospheric CO₂ to produce lower CO₂ and cooler climate. They recognize but do not satisfactorily answer the conundrum that an enhanced global weathering flux to the oceans requires an additional source of carbon dioxide to the atmosphere. Although they propose a lessening of carbon burial as one process (and potential others) that "would have partially counteracted a drawdown of atmospheric CO₂", a partial balance is nowhere near good enough. For example, were today's sinks of atmospheric CO₂ only 10% larger than the sources, CO₂ would be completely removed from the atmosphere-ocean system in a mere 5 Myr.

Perhaps Raymo and Ruddiman intend the net imbalance from the simul-

taneous, rather large perturbations of sources and sinks to be slight enough to weave reasonably-bounded CO₂ changes over 40 Myr? To my mind this would be threading an improbable needle. Although the time history of the Tibetan uplift is not available, Raymo and Ruddiman need to demonstrate that their concept in principle — using a model with explicit mass fluxes and reservoirs, and with numbers they cite, such as a 40% increase in weathering — does not lead to catastrophic drops in CO₂. I suspect that they will be forced to formulate some direct relation between atmospheric CO₂ and weathering, as used ubiquitously in the quantitative system models. Such models routinely produce outputs of temperature and isotopes for comparing to the geological record. I look forward to seeing actual numerical output produced by the Tibetan plateau hypothesis in the context of a dynamic system model.

Tyler Volk

*Earth Systems Group,
Department of Applied Science,
New York University,
New York,
New York 10003, USA*

SIR — Raymo and Ruddiman¹ propose a false dichotomy between climate hypotheses in which variations in atmospheric CO₂ are driven either by accelerated weathering accompanying tectonic uplift or by changes in CO₂ outgassing. Furthermore, they propose that mountain uplift is the primary control on the long-term global chemical weathering rate, and they discount the role of the CO₂ dependence of terrestrial silicate weathering as a major feedback control on atmospheric CO₂ content. We believe that their paper confuses a number of important issues.

Higher atmospheric pCO₂ may lead to higher temperatures, increased runoff, enhanced soil microbial activity and greater area coverage by vegetation at high latitudes, all tending to accelerate mineral dissolution rates. Hence, silicate rock should weather more easily when the atmospheric CO₂ concentration is higher. Because silicate weathering consumes CO₂, this would produce a negative feedback that could modulate long term atmospheric CO₂ content^{2,3}. The calcium and magnesium released to rivers by the chemical weathering of silicate rocks, after entering the ocean and reacting with seafloor basalt, are predominantly buried as carbonates. Hence, on long timescales (>1 Myr), the chemical weathering of silicate rocks must roughly balance the flux of carbon available to be buried as carbonate. In the absence of variations in CO₂ outgassing and for a constant organic carbon sub-cycle, factors such as tectonic uplift, if

it allows silicate rocks to weather more easily, would tend to decrease atmospheric pCO₂ without having any long-term effect on the global chemical weathering rate. To increase the long-term rate, an additional source of carbon is required.

Raymo and Ruddiman propose that the primary source of additional carbon leading to enhanced weathering rates in the mid-to-late Cenozoic is diminished organic carbon burial, which is equivalent to net organic carbon oxidation. (Contrary to Raymo and Ruddiman's characterization, the use of organic burial in carbon cycle modelling is not new⁴⁻⁷.) However, the timing of changes in the Cenozoic carbon isotope record⁸ do not closely correspond to changes in the strontium isotopic record. If terrestrial silicate-weathering rates were to have steadily increased by 40% over the past 40 Myr as proposed by Raymo and Ruddiman, then the net oxidation of this amount of organic carbon would require that O₂ levels were about twice as great as today 40 Myr ago. There is no independent evidence for this extreme situation. Enhanced burial of sedimentary pyrite is the primary possible non-atmospheric source of O₂ for net organic carbon oxidation. However, sulphur isotopic evidence points to little change and possibly some decrease in late Cenozoic pyrite burial rates^{9,10}.

The isotope evidence that Raymo and Ruddiman cite could be explained by processes they did not consider. For example, the carbon isotope signal after 15 Myr may not only reflect diminished organic carbon burial, but rather could be partly a consequence of diminished biological carbon isotopic fractionation^{11,12}. (If the late Cenozoic trend has been towards lower pCO₂, then plants from later in this period should show less selectivity towards lighter carbon isotopes than did plants from earlier.) Also, the mid-to-late Cenozoic strontium isotope signal could be the result of the liberation of radiogenic strontium from old silicates during collisional orogenesis¹³, rather than from an increase in overall weathering rates.

Organic matter burial cannot act as the feedback on atmospheric pCO₂ because factors other than atmospheric pCO₂ (for example, nutrient supply to the oceans) control the rate of organic matter burial¹⁴. The other two possible feedbacks suggested by Raymo and Ruddiman also have serious deficiencies. First, there is more evidence for authigenic calcium silicate formation in marine environments before 40 Myr ago than over the past 40 Myr (ref. 15), indicating that enhanced weathering could not have been balanced by enhanced authigenic calcium silicate formation in the late Cenozoic. Second, cal-

culations indicate that seafloor basalt weathering rates, near the pH of sea water, should be very insensitive to changes in atmospheric CO₂, precluding this process as an important CO₂ feedback mechanism¹⁶. Hence, in the Raymo and Ruddiman model, atmospheric CO₂ could vary widely because the CO₂ concentration is not fixed by a set of feedbacks.

A more inclusive model for the evolution of atmospheric CO₂ should attempt to consider quantitatively all major processes affecting the carbon cycle, including the role of plant evolution as it affects weathering¹⁷, changes in the position of continents relative to climate zones¹⁸, the importance of marine versus terrestrial burial of organic matter¹⁹, the transfer of carbonate from the shelves to the deep sea²⁰, the effect of uplift on weathering¹ and so on. These factors are already considered in a preliminary manner in an existing model⁷.

Raymo and Ruddiman have drawn attention to one potentially important factor in long-term climate change. Still, there is no need to propose a dichotomy between tectonic uplift and CO₂ degassing, because no one factor can explain atmospheric CO₂ changes over all time. Significant changes could well result from a simultaneous combination of any number of forcing factors.

Ken Caldeira

Michael A. Arthur

*Earth System Science Center &
Department of Geosciences,
Pennsylvania State University,
University Park,
Pennsylvania 16802, USA*

Robert A. Berner

Antonio C. Lasaga

*Department of Geology and Geophysics,
Yale University,
New Haven,
Connecticut 06511, USA*

unweathered rock; uplift induced precipitation and runoff). Likewise, proxy records^{21,25–28} indicate an increase in global chemical weathering rates over the past 100 Myr, a time of global cooling.

Both Volk and Caldeira *et al.* regard our challenge to the earlier view of chemical weathering as “conceptual” in nature, and Volk states that our views should be held up to the quantitative scrutiny that more detailed (“matter-conserving”) models can provide. But their criticisms miss a key point implicit in our paper. Such models can only be used for verification if their key parameterizations and assumptions are sound. Matter-conserving numerical models supporting our view have been built^{22,29}. We do not consider the fact that these models work to be “an essential test” of our hypothesis. Rather, a numerical model has been constructed which provides fewer conflicts with existing geological data (the true test).

A second key issue is the need for carbon fluxes sensitive to CO₂ levels and/or climate to provide a negative feedback that avoids a runaway ‘ice-house’ or ‘greenhouse’. Instead of the temperature-weathering feedback, we point to the likelihood of carbon fluxes from other components of the global carbon cycle, such as the organic carbon subcycle¹, basalt weathering²², or some place as yet unforeseen, playing such a role. Contrary to the specific criticisms of Volk and Caldeira *et al.*, small variations in the fraction of carbon buried as organic matter versus carbonate (from ~25 to ~20%) can provide enough CO₂ to balance a 40% increase in the chemical weathering of silicates, carbonates and organic-rich rocks. To oxidize this much carbon would require oxygen levels within ~10% of Berner’s own

estimates of Eocene O₂ levels³⁰. With the exception of a small downturn over the past 10 Myr, sulphur isotopes have been increasing for the past 100 Myr (refs 9, 10, 31) making this postulated oxygen consumption even less of a problem. To attribute the bulk carbonate δ¹³C data of Shackleton to changing photosynthetic fractionation factors would require a decrease much larger than suggested by data (>10‰ over 20 Myr, compared to 5–7‰ change observed over the past 100 Myr; refs 32, 33). And, contrary to the statement of Caldeira *et al.* we did reference previous attempts to model the organic carbon cycle (including refs 5, 7) and attributed our divergence of opinion to the use of different assumptions and data sets (see ref. 34 versus ref. 8).

The Walker and Berner carbon-cycle models revolutionized our thinking of long-term climate evolution and the global carbon cycle. However, mounting evidence suggests a world more complicated than this generation of models would imply. We are sure that our critics would agree that science advances by close scrutiny of both models and data. In this case, we believe sufficient modern (and palaeo) data exist to raise serious questions about the chemical weathering parameterization of the BLAG model (and similar models) of long-term climate change.

Maureen E. Raymo

*Department of Earth, Atmospheric, and
Planetary Sciences,
Massachusetts Institute of Technology,
Cambridge,
Massachusetts 02139, USA*

William Ruddiman

*Department of Environmental Sciences,
University of Virginia,
Charlottesville,
Virginia 22903, USA*

RAYMO AND RUDDIMAN REPLY — The “dichotomy” we really present is between carbon cycle models assuming a positive dependence of silicate and carbonate weathering rates on atmospheric CO₂ and temperature^{2–5,7} and those that do not^{1,21,22}. In all BLAG-type models, including Berner’s⁷, chemical weathering rates decrease as global temperatures fall (and vice versa). This is the result of a dimensionless feedback function included in these models. Viewed in isolation from other factors, this temperature-weathering feedback is of course correct. But we have suggested that an entirely different factor (tectonic uplift) is a more important control of chemical weathering in the real world. Modern river studies^{23,24} support our view that chemical weathering is highly sensitive to uplift-related factors (exposure of folded, faulted, and largely

1. Raymo, M. E. & Ruddiman, W. F. *Nature* **359**, 117–122 (1992).
2. Walker, J. C. G., Hays, P. B. & Kasting, J. F. *J. geophys. Res.* **86**, 9976–9982 (1981).
3. Berner, R. A., Lasaga, A. C. & Garrels, R. M. *Am. J. Sci.* **283**, 641–683 (1983).
4. Garrels, R. M. & Lerman, A. *Am. J. Sci.* **284**, 989–1007 (1984).
5. Lasaga, A. C., Berner, R. A. & Garrels, R. M. in *The Carbon Cycle and Atmospheric CO₂: Natural Variation Archean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 397–410 (Am. Geophys. Un., Washington DC, 1985).
6. Kump, L. R. & Garrels, R. M. *Am. J. Sci.* **286**, 337–360 (1986).
7. Berner, R. A. *Am. J. Sci.* **291**, 339–376 (1991).
8. Shackleton, N. K. in *Marine Petroleum Source Rocks* (eds Brooks, J. & Fleet, A. J.) 423–434 (Geol. Soc. London, 1987).
9. Burdett, J. W., Arthur, M. A. & Richardson, M. *Earth planet. Sci. Lett.* **94**, 189–198 (1989).
10. Claypool, G. E., Holser, W. T., Kaplan, I. R., Sakai, H. & Zak, I. *Chem. Geol.* **28**, 199–260 (1980).
11. Arthur, M. A., Dean, W. A. & Claypool, G. F. *Nature* **315**, 216–218 (1985).
12. Kump, L. R. *Am. J. Sci.* **289**, 390–410 (1989).
13. Palmer, M. R. & Edmond, J. M. *Geochim. cosmochim. Acta* **56**, 2099–2111 (1992).
14. Broecker, W. S. & Peng, T. H. *Tracers in the Sea* (Elsevier, Palisades, New York, 1982).
15. Kastner, M. in *The Sea* **7** (ed. Emiliani, C.) 915–980 (Wiley, New York, 1981).
16. Caldeira, K. *Am. J. Sci.* (in the press).
17. Volk, T. *Geology* **17**, 107–110 (1989).
18. Tardy, Y., Roger, N. K., Probst, J. D. *Am. J. Sci.* **289**, 455–483 (1989).
19. Kump, L. R. *Nature* **335**, 152–154 (1988).
20. Caldeira, K. *Nature* **357**, 578–591 (1992).
21. Raymo, M. E., Ruddiman, W. F. & Froelich, P. N. *Geology* **16**, 649–653 (1988).
22. François, L. M. & Walker, J. C. G. *Am. J. Sci.* **292**, 81–135 (1992).
23. Gibbs, R. J. *Science* **156**, 1734–1737 (1967).
24. Stallard, R. F. in *Physical and Chemical Weathering in Geochemical Cycles* (eds Lerman, A. & Meybeck, M.) 225–246 (Kluwer, Dordrecht, 1988).
25. Shemesh, A., Mortlock, R. A. & Froelich, P. N. *Paleoceanography* **4**, 221–234 (1989).
26. Richter, F. M., Rowley, D. B. & DePaolo, D. J. *Earth planet. Sci. Lett.* **109**, 11–23 (1992).
27. Hodell, D. A., Mead, G. A. & Mueller, P. A. *Chem. Geol. (Isotop. Geosci. Sec.)* **80**, 291–307 (1990).
28. Krishnaswami, S., Trivedi, J. R., Sarin, M. M., Ramesh, R. & Sharma, K. K. *Earth planet. Sci. Lett.* **109**, 243–253 (1992).
29. François, L. M., Walker, J. C. G. & Opdyke, B. AGU-IUGG Spec. Vol. (in the press).
30. Berner, R. & Canfield, D. *Am. J. Sci.* **289**, 333–361 (1989).
31. Lasaga, A. C. *Am. J. Sci.* **289**, 411–435 (1989).
32. Arthur, M. A., Dean, W. E. & Claypool, G. E. *Nature* **315**, 216–218 (1985).
33. Popp, B. N., Takigiku, R., Hayes, J. M., Louda, J. W. & Baker, E. W. *Am. J. Sci.* **289**, 436–454 (1989).
34. Lindh, T. B. thesis (Univ. Miami, 1983).

Political pains

Daniel S. Greenberg

The Mortal Presidency: Illness and Anguish in the White House. By Robert E. Gilbert. *Basic Books*: 1992. Pp. 314. \$25.

PRESIDENTIAL health is a deep concern of the electorate in health-obsessed America — so deep, in fact, that incumbents and candidates frequently lie about their infirmities, often in connivance with their personal physicians. The appearance of ruddy good health, symbolized by sweat-soaked jogging or at least an afternoon of golf, is a main ingredient of political success at the presidential level. But the facts of presidential health often belie the appearances.

While under treatment for chronic, debilitating Addison's disease, John F. Kennedy vigorously denied that he suffered from the affliction. His doctors fended off the press with a variety of obfuscations. Franklin Roosevelt, in terminal decline after 12 years in office, concealed his physical decrepitude, won reelection and died 6 months later. Dwight Eisenhower suffered a heart attack and a stroke and underwent abdominal surgery during his two terms in the White House. Often weak and unable to function in the aftermath of these tribulations, Eisenhower was nonetheless depicted by his physicians and public-relations managers as the beneficiary of extraordinary recuperative powers — and the public believed them. In a footnote to the 1992 presidential election, former senator Paul Tsongas, diagnosed in late November with a recurrence of cancer, admitted that during his unsuccessful run for the Democratic nomination, he had not been fully candid about his health.

The rigours of the office, presidential infirmity and medical deception are the subject of *The Mortal Presidency*. Robert E. Gilbert, professor of political science at Northeastern University, turns in an excellent performance in demonstrating that health and politics American-style are an uncertain and troublesome combination. But along with many others who have delved into the elusive problems of anticipating and coping with presidential frailties, Gilbert is at a loss to offer feasible remedies beyond measures now in place.

The bulk of the book is devoted to the health histories of five presidents of this century — Coolidge, who fell into depression after the death of his young

son; Roosevelt, Eisenhower and Kennedy, with their aforementioned ailments, plus others; and Reagan, who survived a gunshot wound and colon cancer, while his handlers followed the well-established practice of assuring the public that the chief was improving at an incredible pace. Gilbert writes that following Reagan's cancer surgery, "[t]he release of photographs was ... con-

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Bill Clinton — dying to be president?

trolled and arranged for maximum political effect. [Presidential Press Secretary] Speakes describes the first post-operative photograph of the President as being 'artfully arranged to conceal the nasogastric tube that had been inserted in Reagan's nose and was held in place by tape.' The photograph showed the First Lady kissing the President, with her face concealing both tape and tube."

By custom and constitutional amendment, the American political system has tried to reduce the uncertainties of presidential disability. Although there is no means of compelling candidates to present themselves to independent medical panels, custom — and the threat of worrisome speculation — assures the public release of detailed medical records and, invariably, clean bills of health from the candidates' physicians. During the presidential campaign, Bill Clinton quickly produced his medical records when the *New York Times* reported that he had been less forthcoming about his medical status than other

candidates of recent times.

Adopted in 1967 to reduce uncertainties in presidential succession, the 25th Amendment to the Constitution has left a good deal of unease about the passage of power in the nation's greatest office, especially in the foggy circumstances of temporary presidential disability. After labouring long on the amendment, Senator Birch Bayh conceded to his colleagues, "I have never pretended ... that it would cover every possible, conceivable contingency that the mind of man could contrive. I have suggested that it is the best thing that we have been able to come up [with], and it is so much better than anything we have ever had before — namely, nothing." Gilbert agrees with that modest appraisal, stating that the amendment "probably does the best that can be expected" and that "tinkering with it would be unwise."

Associated Press

While health is a sensitive issue in American presidential politics, it is by no means certain that unrestrained medical candour would improve the selection process. Some very sickly people, Kennedy among them, have successfully managed the rigours of the presidency. Furthermore, the predictive powers of medicine do not reliably extend to political performance. What is the electorate to conclude if informed that a candidate suffers from this or that dreadful-sounding disease? When doctors disagree, how can voters decide?

In his principal venture into reform, Gilbert observes that insufficient attention is paid to the mental health of presidents. To correct this, he suggests that "a small mental health unit might be established in every administration as part of the White House medical office to assist the White House physician in diagnosing and treating psychological problems."

The medical desirability of this innovation is indisputable — as is the political impossibility. It should be recalled that in 1972 Senator Thomas Eagleton, candidate for vice-president, was dropped from the Democratic ticket after it was disclosed that he had been treated for depression. In the 1988 presidential election, Michael Dukakis, the Democratic candidate, went to great lengths to dispel scurrilous rumours that he had once consulted a psychiatrist.

When it comes to the presidency, Americans apparently feel that it is better to be disturbed than treated. □

Daniel S. Greenberg is editor and publisher of *Science and Government Report*, 3736 Kanawha Street NW, Washington, DC, 20015, USA.

Madding masses

Hugh Freeman

Crowds, Psychology, and Politics, 1871–1899. By Jaap van Ginneken. Cambridge University Press: 1992. Pp. 269. £35, \$59.95.

MENTION the psychology of crowds to almost anyone with a passing knowledge of the subject, and the automatic response is "Le Bon". Nearly a century after its publication, his *Psychology of Crowds* remains a principal reference point for the study of mass psychology. But since both its text and the author's life have already been thoroughly explored, why the present book?

Van Ginneken explains that over the course of 20 years, he came to realize that the conventional history of crowd psychology was substantially wrong — biased particularly by undue emphasis on French contributions, on those of a few well-known figures, and on a mono-rather than multidisciplinary perspective. He therefore set out to reconstruct how the early modern theories of crowd behaviour evolved, and to interpret their meaning in a culturally relative sense. He has retraced this intellectual tradition back to the work of Taine — seen as its real founder — and rediscovered a missing link (Fournial) with the better known Sighele, Le Bon and Tarde. Each of these basic theorists is examined in terms of his life and his work's social context, intellectual background and influence on science and public attitudes. This is an ambitious project, and one that is broadly successful.

Taine was a polymath; starting as a literary scholar, he became a professor of art history, and then founded scientific psychology in France, almost certainly influencing Freud. On the basis of Darwinism, he saw the excesses of crowds as a retreat from civilization, a view that he formalized in his theory of "dissolution"; experience of the Paris Commune then made him focus on mob rule, which figured prominently in his massive *Origins of Contemporary France*. Van Ginneken acknowledges this pioneering application of social and psychological mechanisms to the study of political processes, but finds Taine uninterested in the reasons why crowds behaved as they did. His political conservatism, further alarmed by the technical and media developments of the *belle époque*, was taken over by Le Bon, who failed to acknowledge this debt, among others.

Because almost all the work on crowd psychology of this period was in France or Italy, it has usually been described as the 'Latin School', but van Ginneken



Rabble rouser — Le Bon's *Psychology of Crowds* was a strong influence on Hitler.

shows that, unlike the French, the Italian writers were essentially criminologists who intervened in the legal consequences of crowd disorders and were politically radical. Sighele saw crowds as developing a kind of mental unity, particularly when they were very large; like others, he moved increasingly to a nationalist position, providing one contribution to the intellectual rag-bag that later made up fascism. Here, the two national trends converged, since Le Bon was also a strong influence on Mussolini, who skilfully manipulated crowd violence to intimidate opponents, as he was on Hitler.

In both Latin countries, physical anthropology was then believed to provide the key to many cultural and social phenomena, including deviant behaviour. (The uniqueness of fingerprints, incidentally, was first reported by an English medical missionary from Tokyo, in a letter to *Nature*). Fournial, a French colonial doctor, is identified by van Ginneken as the missing link in the transition from this scientific blind alley to a more psychological analysis. His chosen mechanism was "social imitation", which shared qualities with hypnosis, and this contribution also was used but unmentioned by Le Bon.

In France, hypnotic suggestion was then a scientific battleground between Charcot's school in Paris, who saw it as necessarily linked to hysteria, and those with a wider view in Nancy; Le Bon used this work freely in interpreting the minds of crowds. Simultaneously with his *Psychology of Crowds*, Breuer and Freud in Vienna were producing *Studies in Hysteria*, which incorporated the same notion of successive levels of consciousness and was influenced by the same debate. The other psychological leg of Le Bon's theoretical structure was in the evolution-dissolution tradition that de-

rived from Darwin and Spencer. The product was pessimistic, authoritarian, élitist and racist, even for its day. Its political background was the abortive Boulanger coup (showing crowds at their most gullible), the first mass socialist demonstrations, and spectacular acts of anarchist terrorism. Through his analysis of behaviour in the mass, Le Bon became advisor to the political élite.

The final theorist, Tarde, appears here largely as a postscript to Le Bon, but he had some importance as founder of the modern study of opinion formation. What emerges clearly from this analysis is that today's reductionism would be very misleading here. As a science, late nineteenth-century psychology was still largely unformed, and the work of its pioneers spread across what are now quite separate disciplines; this may well be what subsequently prevented the study of crowds from becoming a major focus of scientific work.

In all these fields, van Ginneken shows an easy mastery both of the literature and of previously unexploited primary sources. Although he is fluent in English, it is clearly not his native tongue, and the book failed to receive the skilful editing it needed to make sure that the quality of writing equalled that of the underlying scholarship. Regrettably, the text is sprinkled with sloppy phrasing, incorrect words and egregious exclamation marks; we are helpfully told that Hitler was "the German Führer" and are recommended to "see review by McGuire", only to find in the references that it is unpublished. Yet in spite of it all, van Ginneken has made a significant contribution to the intellectual history of crowd psychology. □

Hugh Freeman is editor of the British Journal of Psychiatry, 21 Montagu Square, London W1H 1RE, UK.

From molecular to vernacular

S. M. Walters

The Concise Oxford Dictionary of Botany. Edited by Michael Allaby. Oxford University Press: 1992. Pp. 442. £18.95, \$35 (hbk); £6.99, \$11.95 (pbk).

DICTIONARIES are not to be judged primarily as literary works, but for their coverage and clarity of definition. It is, however, permissible to comment more generally on the preface, for here the aims and intentions of the publishers and the editors are revealed. Michael Allaby's preface is clear, both as to the intended readership — "no longer confined to a minority of specialists" — and style — "a compromise between the two extremes" of encyclopaedia and definitional dictionary. As the proof of the pudding is in the eating, some sampling seems called for.

On the inclusion and definition of individual families and genera, the dictionary scores highly. Of course, selection is a problem, and it would be easy to criticize particular inclusions (35 lines devoted to Haloragidaceae and the genus *Haloragis* may seem excessive to some), but the coverage of all plant groups, including the Lower Plants (not, incidentally, to be found as an entry) seems fairly generous. What of vernacular names? Here lies an insuperable difficulty, although any policy that excluded, for example, 'lychee', 'pink' or 'shield fern' would be unfortunate. But what about 'campion' and 'eye-bright', both widely used English names capable of reasonable definition, but not included?

A more serious test is the coverage of cell and molecular biology. Whether we like it or not, most students of botany are now being taught in departments of plant science or biology, and are exposed to DNA, RNA and the like from the start. Acronyms abound, and 'cistrons', 'codons' and other conceptual *novitates* proliferate. Dictionaries of botany would not sell if they excluded these new terms, but an old reviewer like myself might reasonably ask where this process ends.

A very different area to cover is biological conservation and concern for the environment. Many younger botanists, who might reasonably become users of this dictionary, 'think globally' in a way their predecessors could not, and terms such as 'biodiversity' and 'biosphere' must find their place here. Even the 'Gaian hypothesis' is included, with a rather careful, noncommittal definition.

Zoological dominance is reflected in the definitions of several terms of biosystematic interest, including the crucial definition of the species itself. It is, *pace* Allaby and the Oxford University Press, just not true that: "In taxonomy, it is applied to one or more groups (populations) of individuals that can interbreed within the group, but that cannot exchange genes with other groups. . .". Neither zoologists nor botanists actually operate the taxonomic game on this definition, although zoologists generally seem to pretend that they do. Botanists, in my experience, are more realistic: for every taxonomic species for which there is any substantial body of relevant biosystematic information, there are a hundred for which the morphological criteria are the only ones. Why do we

continue to delude ourselves? The definition of 'deme' — a term originally defined by two botanists — is sufficient to make John Gilmour (one of the two) cry out from the grave, for what is offered as a definition is properly applied to 'gamodeme', and is a zoological aberration.

Dictionaries, finally, are places where the compiler can have a little quiet fun. Turn to page 238, and the running head is "Loppium et chippium". Look down the page and sure enough there it is, coolly defined as "the bark, branches, rotten wood, and leaves from a felled tree". One lives and learns! □

S. M. Walters is at Inland Close, 46 Mill Way, Grantchester, Cambridge CB3 9NB, UK.

M-words and the brain

Susan Greenfield

The Making of Memory: From Molecules to Mind. By Steven Rose. Bantam: 1992. Pp. 355. £16.99.

M-WORDS are particularly prevalent in neuroscience titles because of the catchy oxymorons and alliteration that can be achieved with permutations of 'molecules', 'mind', 'matter', 'make', 'machinery/mechanics' and the like. Steven Rose's book follows in this tradition and introduces a further M-word to savour, memory. There are three central themes: "To chronicle an adventure in research, illustrate the nature of doing science and reflect on the theory of mind".

Rose's account of his own personal adventure into the biochemical basis of memory traces the way in which collective scientific opinion went off the rails, driven by the assumption that memories could be attributed to molecules *per se*. And thus we learn the importance of neuronal organization and canter through the familiar reductionist territory of *Aplysia* and hippocampal long-term potentiation in the sobering light of their questionable physiological relevance to real-life long-term memory. Unfortunately, the more enduring, and thus highly relevant, changes induced by kindling are not discussed at all. Rose moves straight on to describe his own work in seminar-level detail and then turns to recounting how a facile interpretation almost prevented him from seeing a deeper truth. Remarkably, this account of 'pure' science is far from indigestible for the general reader, as Rose is consummate at side-stepping or translating the idiosyncratic terminology and concepts of biochemistry and physiology. The plethora of human interest

anecdotes and the highly personalized narrative more than compensate for the sprinkling of insipid illustrations. The biggest frustration, however, is that the author does not follow up the deeper truth: that the memory for a certain object is distributed in different brain regions. This idea already has an established precedent in the visual system, again not mentioned, and it would have been even more rewarding for the reader for the concept to be set in the general context of the problem of the unifying nature of mind, which is after all within the chosen remit.

The coverage of the day-to-day life of a scientist is highly original and timely. The 'typical' day of a research worker offers a vivid alternative to the image of the white-coated zombie. But the narrative peters out somewhat after a detailed account of the experiments in the morning, just when the reader has acquired a taste for knowing more about the personalities involved and how they interact. Similarly, the journalistic potential in describing conferences and the trials of publication is far from fully tapped. This type of exposé of the human face of scientific research could arguably be a book in itself, especially if ethical considerations are also included. Rose should be strongly commended for confronting the issue of vivisection and discussing it in a frank and open way: he describes exactly what he does to chicks and then offers his reasons. For this initiative he might be forgiven for not telling us what happens in his laboratory in the afternoon.

The third theme relates to mind, a subject that few would deny is now a growth industry. Rose's justification for taking memory as both a title and his

life's work is that it can be seen as the Rosetta stone for matching up mind and brain. The two routes by which he then uses memory to approach mind are those of computer emulation and disruption caused by disease. We take another whistle-stop tour exploring the enduring obsession with modelling the brain using artificial intelligence, and subsequently arrive up against the objections of John Searle (on operational grounds), Gerald Edelman (claiming that the system is ceaselessly dynamic) and Roger Penrose (arguing for a quantum-mechanical and thus unpredictable process). Rose cites the right arguments, but partly for the wrong reasons: no one would claim nowadays that the requisite of dynamism would preclude computers from being like the brain or that they could easily be complex enough to be unpredictable. The really awkward and probing questions raised by quantum mechanics are whether brains function algorithmically and whether the concept of delocalization is applicable to our ultimate understanding of neuronal processes. In view of the distributed properties Rose has discovered about memory, it is surprising that he does not even touch on these concepts. Moreover, he fails even to describe yet another strong indication for distributed memory when considering dysfunction: in W. Penfield's experiments, different memories could be generated by stimulation of the same locus in conscious patients, whereas stimulation at different sites could evoke the same memory.

We have three topics then: the practicalities, results and significance of research, topics that are presented to the reader as parallel rather than complementary themes. Each of the topics throws up a welter of questions that are not followed through because their respective chapters are in the main interspersed, so that attention is switched back and forth without any convincing attempt to draw the threads together, or actually use the Rosetta stone. Rose's gospel of a restless, holistic brain is attractive yet vague, albeit with aspects strongly redolent of Daniel Dennett's "multiple drafts" and Edelman's concept of "re-entry". But the erratic shuttling between the personal, scientific and philosophic conspires against giving the author the breathing space to develop truly new insights. On the other hand, the pace and readability of the text offer a unique smorgasbord in starter-neuroscience, but one that would whet the appetite rather than gratify the palate of those with a taste for M-words and the brain. □

Susan Greenfield is in the Department of Pharmacology, University of Oxford, Mansfield Road, Oxford OX1 3QT, UK.

Making sense

Michael F. Land

Sensory Ecology: How Organisms Acquire and Respond to Information.

By David B. Dusenbery. W. H. Freeman: 1992. Pp. 558. £27.95, \$34.95.

THE study of the senses has always occupied a shifting zone of overlap between physics, physiology and the behavioural sciences of psychology and ethology. Until recently, most of what we knew about the senses of animals came under the general heading of sensory physiology, and was brought together in, for example, the monumental Springer series *Handbook of Sensory Physiology* completed in the early 1980s. The first book I know of called "Sensory Ecology" emerged from a NATO workshop held in 1977 (*Sensory Ecology*, ed. M. A. Ali, Plenum, 1978) and established the agenda for the 'new' subject. The emphasis now is less on the physiology of the sense organs themselves than on the scope and nature of the information 'out there' in the world, and on how animals use this information in the generation and guidance of their behaviour: sense organs are no longer seen as primary, determining what information an animal can take in, but secondary, having developed and been modified by evolution to effect a match between an animal's informational environment and its life-style. Dusenbery takes this general approach, dealing hardly at all with sensory physiology. There are numerous other books called "Animal Senses" and the like that also do this to various degrees, but what makes this book special, and extremely valuable, is the way that the author has brought together the physical underpinnings of the subject, while managing not to be at all dull. There is no other compilation to rival this book, and one would need to work hard to find this information elsewhere.

The book has four parts. An introductory section deals with information and signal processing in general terms; the second discusses stimulus properties in the classical sensory modalities; another deals with stimulus production for the detection of objects via 'active' senses such as echolocation, and also for communication; and a final section is concerned with spatial behaviours such as search and navigation. The first parts tend to be more textbookish, with plenty of equations, tables and generally lucid explanations of physical principles, whereas the later sections are more entertaining for the biologist, with a wealth of intriguing examples of impressive feats of behaviour, ranging from the way

shearwaters hitch around the Pacific on the prevailing winds, to the strategy used by desert ants to locate their nests after foraging trips. The author has certainly hit on a formula that allows both breadth and depth of coverage.

There are some minor faults. A careful reading of Section 3-2 on measuring information reveals an author's name spelt incorrectly, some wrongly printed exponents in Table 3-2, and, amazingly, no mention of C. E. Shannon and W. Weaver's seminal book *The Mathematical Theory of Communication* (1949). But in the section on light, which is my field, I found little to object to. And in a book of this scope the coverage is inevitably uneven. Diffusion and flow, and search strategies associated with them, perhaps get more than their fair share of treatment, but this is the author's subject and some indulgence is proper. No major topic seems to be skimmed, however. Another problem lies in knowing where the boundaries of the subject are. In the section on light, for example, the author starts to stray into questions raised by visual illusions and by *Gestalt* psychology. Although the information an animal seeks does indeed depend on the meanings that the animal can attach to it, the psychology of the senses is such a vast topic in its own right that it might be as well for the fledgling science of sensory ecology not to claim too much of it — for a while anyway!

This is a work of real stature. By bringing so much otherwise difficult material together in a user-friendly way, the book may itself give research in the area a boost. Researchers, students and teachers of every discipline that deals with the senses will want a copy, and every university library should have at least one. I think it is a little too tough as a text for biology undergraduates, who have a traditional terror of equations, but it is certainly an important work of reference for both advanced undergraduate and postgraduate courses. □

Michael F. Land is in the School of Biological Sciences, University of Sussex, Brighton BN1 9QG, UK.

New editions

■ *A Genetic Switch* by Mark Ptashne (2nd edn). Refines what is known about gene regulation in phage lambda and includes two new chapters that extend the treatment to higher organisms. Blackwell Scientific/Cell Press, £17.50 (pbk). For a review by Sydney Brenner of the first edition, see *Nature* **324**, 179 (1986).

■ *Animal Cell Culture: A Practical Approach* edited by R. I. Freshney (2nd edn). IRL (Oxford University Press), £19.50 (pbk).

Agnathans and the origin of jawed vertebrates

Peter Forey & Philippe Janvier

The origins of jawed vertebrates (gnathostomes) lie somewhere within the ranks of long-extinct jawless fishes, represented today as the lampreys and hagfishes. Recent discoveries of hitherto unknown kinds of jawless fishes (agnathans), together with re-examination of known agnathans and advances in systematic methods, have revitalized debates about the relationships of ancient fishes and given fresh insights into early vertebrate history.

FOR many years fossil agnathans were known primarily by cephalaspids (Osteostraci) (Fig. 1d) and pterapsids (Heterostraci) (Fig. 1c), accompanied by a small group of anaspids (Anaspida) (Fig. 1f) and the very poorly known thelodonts (Fig. 1l). All of these fishes occur in Silurian and Devonian deposits and were most numerous in the late Silurian/early Devonian (about 420–390 million years before present (Myr BP)). These extinct fishes were unlike their modern counterparts in possessing a dermal skeleton of minute scales and/or bony plates. They are therefore sometimes known collectively as ostracoderms (bony-skinned) and it is thought that one or other ostracoderm groups gave rise to the gnathostomes.

During the past 15 years, important discoveries of new kinds of extinct agnathans have been made worldwide which considerably increase our knowledge of the variety of agnathans and, at the same time, question some of the traditional ideas about the interrelationships of these early fishes and the history of vertebrate specializations. Some of the important finds are listed.

New fossil finds

(1) In 1977 *Arandaspis* was described from the Lower Ordovician of Australia¹ and in 1986 a very similar form, *Sacambaspis*, was described² from the Upper Ordovician of Bolivia (Fig. 1h). Both of these are early representatives of articulated heterostracan fishes with well developed dermal skeletons and they may offer insights into the anatomy of the earliest vertebrate armour.

(2) From the Silurian and Devonian of China and northern Vietnam we now have evidence^{3–5} of a highly distinctive and varied group of agnathans known as galeaspids (Galeaspida) (Fig. 1e). The internal structures of some of these galeaspids are beautifully preserved⁶ such that it is possible to make detailed anatomical comparisons with cephalaspids and gnathostomes.

(3) Very recent discoveries⁷ of articulated thelodonts from the Silurian and Devonian of northwestern Canada have revealed a surprising new body shape among agnathans and traces of a stomach, as well as providing evidence of the affinities of this enigmatic group.

(4) Another new agnathan body form was revealed with the recent discovery of pituriaspids⁸ from the Middle Devonian of Australia (Fig. 1g). Superficially, they resemble the galeaspids but as yet very little is known about these animals and any statement about their affinities must be speculative.

(5) Last, among the recent fossil finds there has been the discovery of a superficially hagfish-like animal which carries a feeding apparatus consisting of conodonts^{9–11}. Conodonts are tiny phosphatic tooth-like structures, commonly occurring in marine rocks ranging from the Cambrian to the Triassic, and of great biostratigraphic importance for correlating rock sequences. Their biological affinities, however, have never been satisfactorily explained. The discovery of this conodont-bearing animal suggests that at least some conodonts may be among the most primitive of vertebrates.

Alongside the discoveries of these new animals there has also been a considerable increase in our knowledge of the anatomy of heterostracans¹² and osteostracans^{13,14}. Many new detailed

comparisons may be made between the fossils and Recent representatives, leading to new theories about the relationships between agnathan groups and new theories about how jawed vertebrates acquired their characters.

Lampreys, hagfishes and relatives

Lampreys (Petromyzontiformes) and hagfishes (Myxiniiformes) are eel-like, naked animals with no trace of dermal armour or paired fins. Both have a single median nostril and circular mouth (accompanied by a sucker in lampreys), a complex toothed tongue operated by circular and longitudinal muscles, and pouch-like gills opening through circular openings. The feeding habits of both are rather specialized. Most lampreys are filter-feeders as young and ectoparasitic blood-suckers as adults, whereas hagfishes are predators or they rasp away the flesh of dead or moribund fishes. Because of these common anatomical and behavioural features lampreys and hagfishes were thought to be most closely related to one another among the modern fauna and were recognized as cyclostomes (Fig. 2A, a).

This theory carries with it the implication that the cyclostome characters mentioned above were not part of the history of gnathostomes but were instead specializations restricted to lampreys and hagfishes.

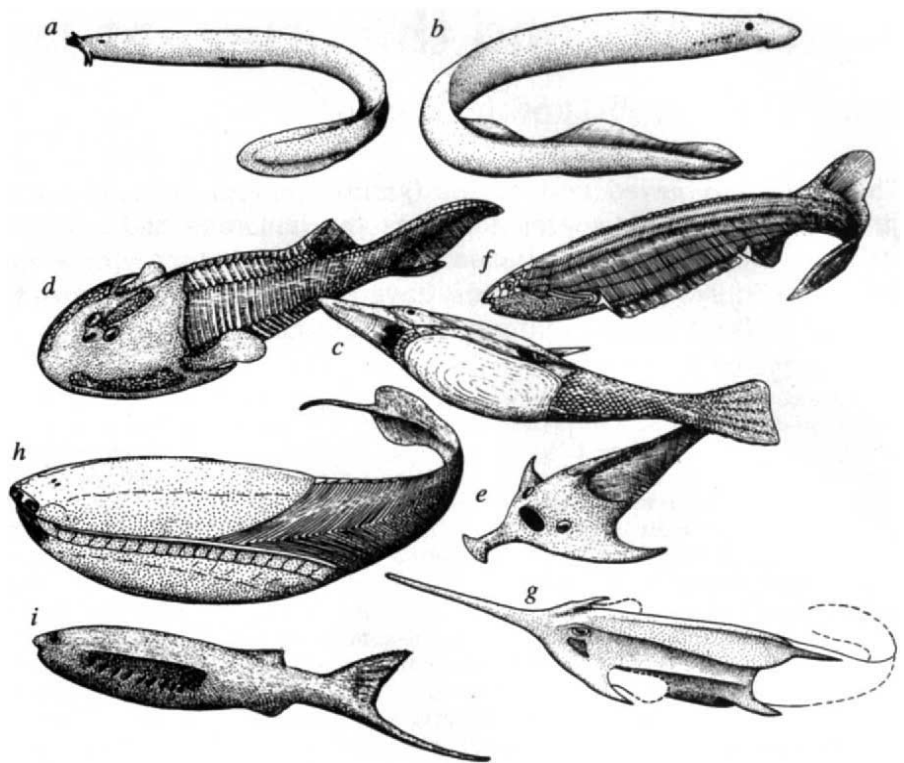
Lampreys and hagfishes have long, separate histories. Fossil representatives of lampreys and hagfishes may be found in the Upper Carboniferous (230 Myr BP) where, in most preserved details, they are similar to their modern counterparts^{15–18}.

Fossil agnathan groups, such as the heterostracans, anaspids and osteostracans were considered related to one or other of the modern types (Fig. 2A, b). Osteostracans (cephalaspids) have long been considered close relatives of lampreys despite a total difference in overall body shape. Like lampreys, cephalaspids show a single nostril joined with a single hypophyseal opening which opens together on top of the head as a keyhole-shaped opening. This is a highly unusual feature, and for many years palaeontologists have interpreted cephalaspids using the lamprey as the model^{19,20}. Anaspids also show a dorsally placed nasohypophyseal opening and are therefore considered to be closely related to lampreys and cephalaspids. The heterostracans were associated with the hagfishes, chiefly because they lacked the specializations of the nasohypophyseal opening and because they have a single branchial opening (see below).

This theory, relating armoured forms to each of the living groups, implies that any feature such as armour, scales or modified fins were independently lost during the history of hagfishes and lampreys.

Another theory concerning the relationships of the Recent taxa suggests that lampreys are more closely related to gnathostomes than either is to hagfishes, implying that hagfishes are the most primitive of craniate animals (Fig. 2B, a). Lampreys and gnathostomes share a great many features which are, at the same time, absent from hagfishes^{21,22}, such as well developed neural and haemal arches (the beginnings of a true backbone) and radial muscles in the fins (which allow voluntary flexing of the fins), true neuromast organs distributed along a lateral line,

FIG. 1 Diversity of jawless fishes. Agnathans are represented today by *a*, hagfishes (*Myxiniiformes*) and *b*, lampreys (*Petromyzontiformes*). Both are naked eel-shaped and without paired fins. They are collectively known as cyclostomes (round-mouths). In the Palaeozoic there were several different groups of agnathans showing a wide variety of morphology of bony dermal armour. These are collectively known as ostracoderms. *c*, Pteraspids (*Heterostraci*) and *d*, cephalaspids (*Osteostraci*) were very common and their solid head shields are often preserved in Upper Silurian and Lower Devonian rocks of North America, Europe and Siberia. The pteraspids (*Pteraspis* illustrated) were the dominant group of heterostracans. Their armour consisted of large dorsal and ventral oval shields surrounded by a series of smaller plates. The eyes were placed laterally and the mouth surrounded by tiny rod-shaped plates. In cephalaspids (*Cephalaspis* illustrated) the eyes were placed dorsally with a single nasohypophysial opening and the flattened head shield was fenestrated by enlarged sensory fields connected to the inner ear. They were enhanced parts of the lateral line system and probably used to detect water vibrations. Cephalaspids had pectoral fins. *e*, Galeaspids (*Galeaspida*) were very unusual agnathans, also with solid head shields, and restricted to eastern Asia. Like cephalaspids, the galeaspid head shield was flattened and in some species may have an unusual shape (*Sanchaspis* illustrated here). The eyes were placed more laterally and there was a large median opening leading to the nasal sacs and the mouth, *f*, Anaspids (*Anaspida*) were a small, poorly known group of fusiform fishes known from North America, Europe and China. They had a dorsal nasohypophysial opening like cephalaspids and lampreys. The gill openings were arranged in an oblique line as in lampreys. Some species, such as *Pharyngolepis*, illustrated here, had long paired fin folds running along the ventral sides of the body. *g*, Pituriaspids from the Devonian of Australia, have only recently been discovered. The shield of *Pituriaspis* was long, sutureless and pierced by triangular openings, of unknown function, on either side of the eyes. It is



probable that paired pectoral fins were present as in cephalaspids. *h*, The earliest known vertebrates (*Arandaspida*) show that by early Ordovician times vertebrate life had already developed complex body armour. *Sacambaspis*, illustrated here, shows large dorsal and ventral head shields separated by tiny branchial plates and a body covered with elongated scales. *i*, The thelodonts (*Thelodontida*) are poorly known, usually found as minute isolated scales. *Thelodus* appears to have been flattened bearing several gill openings beneath pectoral flaps. Some newly discovered thelodonts suggest that, at least some, were markedly compressed laterally.

extrinsic eye muscles, nervous regulation of the heart and hyperosmoregulation.

This theory raises the possibility that cyclostome characters such as the median nostril, complex tongue and pouch-like gills are either primitive craniate characters and are truly part of the history of gnathostomes, or are convergencies, that is, accidental similarities, developed independently in lampreys and in hagfishes.

There have been several suggestions as to how fossil agnathans

may fit into a scheme such as this. Usually, one or other fossil group is considered more closely related to gnathostomes than to any other agnathan, living or fossil. One theory is shown in Fig. 2*B, b*, another in Fig. 4. With these different theories come different conclusions about the history of the skeleton, fins, nasal sacs and so on.

Choice between these competing theories of relationship among the Recent forms has not received consensus (compare refs 23, 24 to refs 21, 22, 25). Unfortunately, molecular evidence,

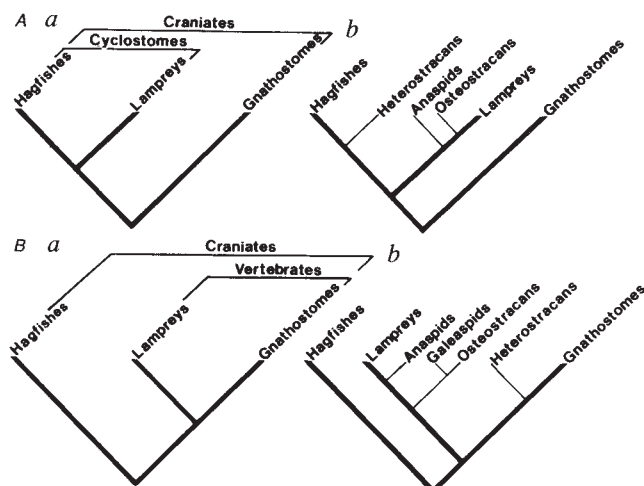


FIG. 2 Theories of craniate relationships. On the left are shown two theories of the relationships amongst modern animals. On the right there are two schemes including fossil agnathan groups congruent with the classifications to the left. *A, a*, A traditional classification which suggests that lampreys and hagfishes are most closely related to one another. *A, b*, This classification⁴⁴ assumes that the fossil groups are related to one or other of the Recent agnathans (galeaspids were not known at this time). *B, a*, This theory suggests that lampreys are more closely related to gnathostomes than to hagfishes (agnathans are considered paraphyletic). *B, b*, This classification^{36,39} is one of several suggesting that one or other of the fossil groups are more closely related to jawed vertebrates than they are to other agnathans. In this case heterostracans are considered the closest relatives of gnathostome; this is based on the speculation that heterostracans had paired nasal sacs and paired nostrils as seen in jawed vertebrates. Thelodonts are not included in these classifications, they are too poorly known.

which has proved itself useful in other areas of disagreement, has yet to prove itself here. Amino-acid data for a variety of proteins has accumulated throughout the 1980s. Sequences for globin molecules are available for lampreys, hagfishes and a variety of gnathostomes, and comparisons of these sequences²⁶ imply that lampreys and hagfishes are sister groups. But this may be misleading because agnathan globin is homologous with both gnathostome myoglobin and haemoglobin. It is probable that the agnathan globin sequences simply represent the primitive craniate sequences from which other gnathostome sequences were derived, perhaps by gene duplication.

Another line of molecular evidence comes from nucleotide sequences of 18s ribosomal RNA²⁷. This analysis used a bootstrapping technique and examined 1,631 base pairs. When cephalochordates or cephalochordates plus tunicates were used as the outgroup, the analysis suggested that lampreys and hagfishes are sister groups. But when tunicates alone were used to root the analysis, the grouping of lampreys and gnathostomes was supported.

A somewhat similar result was obtained using 28S rRNA. The results of analyses are equivocal, and these appear to depend on which outgroup (cephalochordate or tunicate) is chosen to root the analysis. Lecointre (unpublished memoir, DEA, University of Paris VII (1989)) has examined sequences of 361 nucleotides of the 5' extremity of 28S rRNA and found that there are only two base-pair positions, out of 361, common to lampreys and hagfishes: and there is only one position uniquely shared by lampreys and gnathostomes. This is hardly conclusive evidence for one theory or another. Therefore, the molecular data, while adding a new dimension, also requires additional explanation and does not clearly arbitrate between the competing theories of relationship. It is clear that one of the challenges of molecular systematics is to deal with sequence data of ancient and possibly highly specialized lineages.

Significance of the new fossil finds

Heterostracans. One of the chief reasons for believing that heterostracans were fossil relatives of hagfishes (Fig. 2A, b) centred on the fact that in the modern hagfish, *Myxine*, and heterostracans there is a single external branchial opening emptying from several internal gill pouches. This is undoubtedly a specialization which developed from the condition where there are many separate branchial openings as seen, for instance, in lampreys, anaspids, cephalaspids, galeaspids and thelodonts as well as in cephalochordates (one likely sister group of craniates). Although *Myxine* shows a single paired branchial opening, another Recent hagfish (*Eptatretus*) has a series of separate openings and this has cast doubt on reasons for associating hagfishes and heterostracans.

With the discovery of *Arandaspis* and *Sacambambaspis* this doubt is reinforced. These earliest armoured vertebrates are clearly related to heterostracans: they have large dorsal and ventral shields encasing the head and a similar pattern of lateral lines penetrating the shields. At the same time they show a series of separate gill openings wedged between these shields and each associated with a tiny branchial plate. A similar series of separate openings has recently been found in *Astraspis*²⁸ which is another early heterostracan but without the large shield-like armour. Thus, these early heterostracans show a primitive vertebrate condition and suggest that the single branchial opening was developed independently within the heterostracan lineage and within the hagfish lineage.

The discovery of articulated Ordovician heterostracans has other more general implications for our ideas of the evolution of the vertebrate skeleton and tail structure.

For nearly a century it was assumed that the primitive vertebrate armour consisted of scales and small polygonal bony plates. These had been found in the fragmentary vertebrates such as *Astraspis* and *Eriptychius*, long known from Middle Ordovician rocks of North America, and they were often regar-

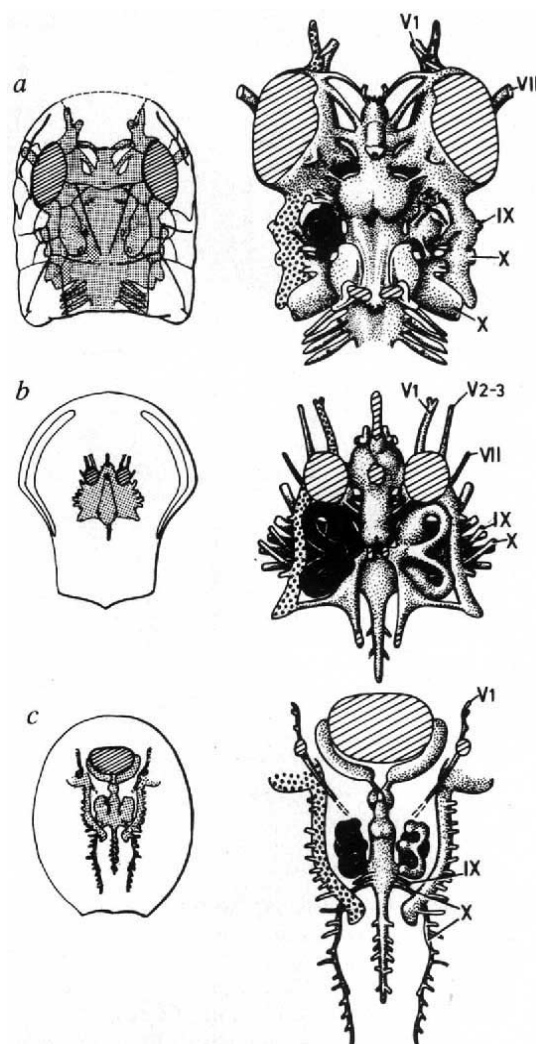


FIG. 3 Internal anatomy. In recent years we have learned much about the internal anatomy of cephalaspids and galeaspids which had a bony skeleton surrounding the brain and gills. There is a remarkable similarity between these fishes and primitive gnathostomes in the structure of the orbit, the labyrinth of the ear and in the position of the head vein. a, *Brindabellaspis*, a placoderm gnathostome (from ref. 45); b, *Norselaspis*, an osteostracan (from ref. 13); c, *Duyunolepis*, a galeaspid (ref. 46). The outline of the head (left) shows the position of the internal cavities (right). The head vein is dotted, the labyrinth of the inner ear is black. Some of the cranial nerves are indicated by roman numerals.

ded as showing primitive vertebrate conditions from which the large shields of more typical pteraspids developed. *Arandaspis* and *Sacambambaspis* both show large continuous head shields and a trunk covered with elongated scales (Fig. 1h). This suggests that considerable diversification of the vertebrate skeleton had already taken place by Lower Ordovician times and, furthermore, that it is not possible on the basis of antiquity alone to predict whether a micromeric or macromeric skeleton is the more primitive vertebrate condition.

The detailed histology of the bone of *Arandaspis* and *Sacambambaspis* is, as yet, unknown. But the Ordovician agnathans *Astraspis* and *Eriptychius* were originally associated with heterostracans because the skeleton was made of aspidin (acellular bone), a condition regarded by some as primitive and by others as derived. The early occurrence of aspidin has always strengthened the opinion that it is the more primitive condition. Recently, however, investigation²⁹ has confirmed earlier suspicions³⁰ that the presence of cellular bone is of equal antiquity. Furthermore the particular type of cellular bone found in the

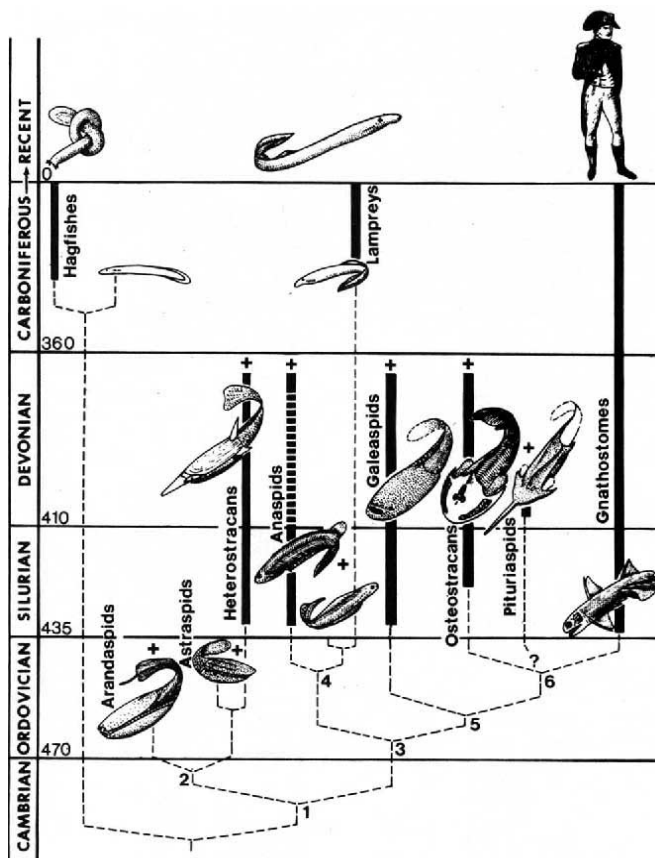


FIG. 4 Our favoured phylogenetic tree of principal agnathan groups with stratigraphic ranges superimposed (bold black lines). This tree is simplified from a PAUP analysis, the details of which will be published elsewhere. The position of pituriaspids and thelodonts is not clear. Some of the skeletal characters supporting the nodes of this classification are as follows: (1) lateral line sensory canal, two semicircular canals in the ear, bone (lost in lampreys). (2) Two large plates (dorsal and ventral) covering the head, separate branchial plates. (3) True dorsal and anal fins with fin rays, asymmetrical tail shape. (4) Branchial openings arranged in a slanting line (suggesting a separation of pharynx from oesophagus as in modern lampreys). (5) Perichondrial bone, large head vein placed dorsally, large orbits. (6) Dermal bone with cells, epicercal tail with modified scale cover (hinge line), paired pectoral fins, open endolymphatic duct, sclerotic ossification.

same deposits as *Eriptychius* suggests that cephalaspids (osteostracans) may be equally as old as heterostracans.

Traditional ideas about the primitive vertebrate tail were derived from studies on pteraspid heterostracans. These ideas^{31,32} proposed that the tail was asymmetrical, with a larger lower lobe and, when in use, tended to drive the head up to compensate for a heavy head shield and a lack of paired fins. *Sacambambaspis* and several other recently discovered well preserved heterostracans from the Silurian of the North West Territories³³ and the Canadian Arctic³⁴ all show perfectly symmetrical tails. The tail was deep with enlarged ridge scales along the upper and lower edges for stiffening, and the fan was strengthened by special radiating rows of enlarged scales (Fig. 1e). This type of tail drove the fish directly forwards.

Cephalaspids. During the last decade or so cephalaspids have been re-examined in great detail^{13,14}. Leaving aside the undoubted similarities in the nasohypophysial opening between cephalaspids and lampreys (see above), these recent studies show that there is a striking resemblance between the cephalaspid skull and the skull of an early jawed vertebrate such as the placoderm (Fig. 3). Externally the cephalaspid body, like that of primitive gnathostomes, shows a heterocercal tail and paired fins (only pectoral fins are developed in cephalaspids). These similarities, together with several other less obvious features, suggest that cephalaspids are more closely related to gnathostomes than either is to lampreys. This would mean that the single nasohypophysial opening is either a convergent feature or a feature primitive for craniates.

Galeaspids. The galeaspids are a group of about 40 species which superficially resemble cephalaspids in having a large, flattened and solid head skeleton which encloses a complex brain and inner ear (Fig. 3). They have been considered as directly related to cephalaspids^{35,36}. But unlike cephalaspids they lack paired fins and, instead of a dorsal nasohypophysial opening, they have a large median dorsal opening which communicates with paired nasal cavities and, further below, with the pharynx. In some respects this arrangement is similar to that seen in recent

hagfishes where a single nostril opens into a duct which passes the single nasal sac before opening into the roof of the mouth. **Thelodonts.** Thelodonts are usually represented in the fossil record as isolated scales which, because of their abundance, are useful for stratigraphic correlation. But we know very little about the animals which bore these scales and although a few articulated remains have been known for several years, they show very little anatomical detail and it is not easy to identify immediate relatives. Hitherto, known thelodonts appear to be flattened dorso-ventrally and to have paired flaps on either side of the head.

Very recently, discoveries of some new and very well preserved thelodonts from the Silurian of British Columbia were announced⁷. These new thelodonts show an unexpected body shape in which the body is deep and compressed from side to side and the tail is perfectly symmetrical, rather like that of early heterostracans. They also show that several very different types of scale could be borne by the same animal, including some complex scales immediately surrounding the gill openings. Interestingly, the gill openings are arranged in an oblique line, a situation very reminiscent of the gill opening of anaspids and lampreys. This may suggest that at least some thelodonts are related to these two groups.

Another interesting feature about these new thelodonts is that there is evidence that a well developed stomach was present. Lampreys and hagfishes lack a differentiated stomach, and it is generally assumed that all agnathans were similarly microphagous with no need for a stomach. Jaws, macrophagy and stomachs were thought to go together. However, in these thelodonts we have evidence that stomachs preceded jaws in vertebrate history.

Conodonts. Conodonts have traditionally been used as form taxa, useful for stratigraphic correlation within Palaeozoic rocks. During the last few years there have been a number of discoveries suggesting that conodonts may be vertebrates.

Briggs and colleagues^{9,10} discovered fossils of 4 cm-long eel-like animals from the Early Carboniferous of Scotland, each of which carried a set of conodont elements at the anterior end. Each of these animals, one of which has been named *Clydagnathus*, shows a laterally flattened body, caudal fin and 'V-shaped' muscle blocks, together with the faint outline of a notochord and large paired eyes. These are features typically found in chordates, although it is difficult to make more precise comparisons.

More recently the histological detail of the minute conodont elements themselves has been re-examined. Most conodonts, or euconodonts, are tooth-like phosphatic structures with a shiny outer layer and an inner core. Recent research³⁷ advocates that the minute spaces within the substance of the core, which have

long been known to exist, are true bone cell or osteocyte cell spaces. The basal tissue of the conodont has also been compared to globular calcified cartilage as seen associated with the earliest authenticated bone of *Eriptychius*. But the surface layer is far more problematic. It appears to be composed of centrifugally deposited lamellar tissue, superficially like vertebrate enamel. But in the conodonts examined so far the crystalline structure within this layer is very variable. In two species examined the crystals lie parallel to the surface and in another two they lie at a steep angle. Neither type corresponds precisely to that seen in vertebrate enamel, and the extreme variation in crystal orientation is puzzling. Furthermore, there appears to be a total absence of dentine, which is unexpected if conodonts are vertebrates. Dentine is universal in vertebrates and is thought to be the most primitive of vertebrate hard tissues²⁹.

Clearly, much more research needs to be done on a wide variety of conodonts before the histological structure may be used as evidence of vertebrate affinities. If that supportive evidence is forthcoming then the implication is that vertebrates existed in the Late Cambrian, 510 million years ago, and some 40 million years before the appearance of vertebrate armour. Conodonts are not considered further here.

Relationships and ideas of vertebrate history

The new features that have been discovered by re-examining Recent taxa alongside the many new fossil finds during the past 15 years lead to classifications that suggest that lampreys are more closely related to gnathostomes than either is to hagfishes. In other words, lampreys and hagfishes do not form a monophyletic group. One of the classifications that inserts the fossil groups is illustrated in Fig. 4. This classification is based on a cladistic analysis of morphological characters using the PAUP parsimony program. Details of this classification will be published elsewhere. Interpreting this as a phylogeny, we can use this classification to trace the history of the transformation of some characters leading to the condition in gnathostomes.

One of the most debated areas of anatomy is the early history of the vertebrate nasal and hypophysial regions already mentioned above. It is not clear whether paired or single nasal sacs and nostrils are primitive for craniates. The phylogeny given here predicts that the condition in hagfishes is primitive. That is, the earliest craniate had a single median inhalent duct conveying water over the unpaired nasal sac on its way to the pharynx and gills. The gnathostome condition, with paired nasal sacs not associated with a through duct, would therefore be a derived condition.

The nature of the nasal sacs in heterostracans has been the subject of some debate. Stensiö³⁸ reconstructed the snout after the hagfish model. Others^{36,39} have suggested that there were paired nasal sacs opening separately and gave rise to the theory that heterostracans might be ancestral to jawed vertebrates (Fig. 2B, b). As long as the endoskeleton of the heterostracans remains unknown there must always be doubt about the pattern. Some faint impressions on the inside of the dermal skeleton suggest that at least paired nasal sacs may have been present⁴⁰ (unlike in modern hagfishes). But the inhalent duct leading to the outside may well have been single. We now know that this was the condition in galeaspid where paired nasal sacs opened into median duct and the duct itself communicated with the pharynx. This may represent an example of a second stage of transformation. Subsequent modification of this pattern would be the confluence of both nasal sacs and the failure of the inhalent duct to communicate with the mouth in lampreys and cephalaspid. A different modification would have taken place in gnathostomes where the separate nasal sacs remained large but the common inhalent duct was lost.

The phylogeny suggested here implies that hagfishes are primitively naked, whereas lampreys are secondarily so. For the history of the dermal skeleton the recent finds have tended to complicate the issue because it is no longer clear from the

stratigraphic record whether cellularity or acellularity is primitive, or whether a solid shield or polygonal plates are primitive.

The evolution of fins may also be traced across this phylogeny. Hagfishes have a weakly developed symmetrical tail and the web is supported by poorly developed cartilage rods without associated muscles. This was probably very similar in the heterostracans except that here the tail lobe was scale covered, perhaps with some special scale rows aligned along the length of fin (see above). From here anaspids, lampreys and perhaps some thelodonts developed a downturned asymmetrical tail (hypocercal), and in the first two there were definite endoskeletal fin rays with, in at least lampreys, associated musculature. In cephalaspid the tail is turned upwards (epicercal) as in gnathostomes and here too there were probably muscularized endoskeletal fin rays developed. The tail of galeaspid remains poorly known.

Paired pectoral fins of gnathostome-type were only developed in cephalaspid (and perhaps pituriaspid) where there is evidence of an endoskeleton and associated muscles. Some anaspids have specialized elongate paired fins which may have been developed and restricted to this group (Fig. 1f).

The outline above suggests that agnathans are a paraphyletic group (that is, some are more closely related to jawed vertebrates than to other agnathans) and therefore the presence of circular mouths and pouch-like gills was part of gnathostome history. The acquisition of many gnathostome characters occurred through several transformation series which can be traced across the phylogeny of agnathans. But some characters appear suddenly, and this applies to the feature which characterizes the gnathostomes: the development of jaws.

The way in which jaws developed remains unknown. There is a long-held belief that jaws developed from a true mandibular gill-bearing arch. Yet no animal showing this has ever been found. The idea stems from the belief that in gnathostomes the mandibular arch is a serial homologue of the hyoid arch and the more posterior gill arches. Agnathans have gill arches but no jaws. Hence jaws might be a transformation of the anterior-most gill skeleton. But we know that the gill lamellae in a lamprey develop medially to the supporting skeleton, whereas the gills of a gnathostome develop laterally to the skeleton. For this reason some workers have denied any homology^{41,42}. One possible alternative may be to suggest that jaws correspond to the velum in a larval lamprey, a special pumping device located at the entrance to the pharynx. The velum develops from the same embryonic segment as the mandibular arch, and part of its skeleton is located more medially than that of subsequent arches.

Summary

The discovery of many new facts and new types of agnathans with hitherto unknown anatomical features has questioned some of our traditional ideas. At the same time, it has permitted the formation of a more precise theory of relationships from which we may deduce evolutionary pathways.

We may ask if we would have different views on the interrelationships of modern craniates if we had no fossils. In this case we doubt it. But if we asked about the history of the appearance of characters then the answer is 'yes'. Fossils do throw light on the history of the lateral line and tail. In these cases the fossils show that these features have developed through an ordered series of transformations leading to conditions in the jawed vertebrates. The fossils show that paired fins constructed like those in gnathostomes are a relatively late development but that there may have been an experiment with elongated paired fins as seen in anaspids. Fossils also show that the structure of the nasohypophysial region is considerably variable and more complicated than the Recent animals might suggest. At the same time, the great variety of skeletal anatomy has complicated our theories of the early evolution of the vertebrate dermal skeleton.

Interesting future directions of study might be attempts to understand the varied locomotory adaptations of agnathans which show different development of paired and median fins. Detailed anatomical studies need to be done alongside modern hydrodynamic experiments⁴³. There is also considerable scope for informed speculation about the feeding mechanisms among

the extinct agnathans. The wealth of new information is a reminder that there is still plenty of scope for new studies in old fishes. □

Peter Forey is at The Natural History Museum, Cromwell Road, London SW7 5BD, UK; Philippe Janvier is at the Muséum National d'Histoire Naturelle, 8 rue Buffon, 75005 Paris, France.

1. Ritchie, A. & Gilbert-Tomlinson, J. *Alcheringa* **1**, 351–368 (1977).
2. Gagnier, P.-Y., Bileck, A. R. M. & Rodrigo, S. *Geobios* **19**, 629–634 (1986).
3. Pan Jiang, K. *Proc. Linn. Soc. N.S.W.* **107**, 303–319 (1984).
4. Janvier, P. *J. vert. Paleont.* **4**, 344–358 (1984).
5. Tong-Dzuy, T. & Janvier, P. *Bull. Mus. nat. Hist. nat. Paris* **12**, 143–223 (1990).
6. Halstead, L. B. *Nature* **282**, 833–836 (1979).
7. Wilson, M. V. H. & Caldwell, M. W. *Nature* (in the press).
8. Young, G. C. in *Early Vertebrates and Related Problems of Evolutionary Biology* (eds Chang, M. M., Liu, Y. H. & Zhang, G. R.) 67–85 (Science Press, Beijing 1991).
9. Briggs, D. E. G., Clarkson, E. N. K. & Aldridge, R. J. *Lethaia* **16**, 1–14 (1983).
10. Aldridge, R. J., Briggs, D. E. G., Clarkson, E. N. K. & Smith, M. P. *Lethaia* **19**, 279–291 (1986).
11. Briggs, D. E. G. *Bull. Field Mus. nat. Hist.* **55**, 11–18 (1984).
12. Bileck, A. *Les Hétérostracés (Vertébrés, Hétérostracés) de l'horizon Vogti (Groupe de Red Bay, Dévonien inférieur de Spitzberg)* (CNRS, Paris, 1985).
13. Janvier, P. *Palaeovertebrata* **11**, 19–130 (1981).
14. Janvier, P. *Les Cephalaspides du Spitzberg* (CNRS, Paris, 1985).
15. Bardack, D. & Zangerl, R. *Science* **196**, 1265–1267 (1962).
16. Janvier, P. & Lund, R. *Geobios* **19**, 647–652 (1986).
17. Bardack, D. *Bull. Soc. Hist. nat. Autun (France)* **116**, 97–99 (1985).
18. Bardack, D. *Science* **254**, 701–703 (1991).
19. Stensiö, E. A. *Skrifter om Svalbard og Ishavet* **12**, 1–391 (1927).
20. Stensiö, E. A. *The Cephalaspids of Great Britain* (British Museum (Natural History), London, 1932).
21. Hardisty, M. W. in *The Biology of lampreys* (eds Hardisty, M. W. & Potter, I. C.) 165–259 (Academic, London, 1982).
22. Janvier, P. *J. vert. Paleont.* **1**, 121–159 (1981).
23. Schaeffer, B. & Thomson, K. S. in *Aspects of Vertebrate History* (ed. Jacobs, L. L.) 19–33 (Museum of Northern Arizona Press, Flagstaff, 1980).
24. Valden, D. W. *Zool. J. Linn. Soc.* **84**, 291–300 (1985).
25. Løvtrup, S. *The Phylogeny of the Vertebrata* (Wiley, London, 1977).
26. Goodman, M., Weiss, M. L. & Czelusniak, J. *Syst. Zool.* **31**, 376–399 (1982).
27. Stock, D. W. & Whitt, G. S. *Science* **257**, 787–789 (1992).
28. Elliott, D. K. *Science* **237**, 190–192 (1987).
29. Smith, M. M. & Hall, B. K. *Biol. Rev.* **65**, 277–373 (1990).
30. Denison, R. H. *Fieldiana Geol.* **16**, 131–192 (1967).
31. Stensiö, E. in *Traité de Paléontologie* Vol. 4 (ed. Piveteau, J.) (Masson, Paris, 1961).
32. Kermack, K. A. *J. exp. Biol.* **20**, 23–27 (1947).
33. Soehn, K. & Wilson, M. V. H. *J. vert. Paleont.* **10**, 405–419 (1990).
34. Dineley, D. L. & Loeffler, E. J. *Spec. Paps. Palaeont.* **18**, 1–214 (1976).
35. Janvier, P. *Bull. Mus. nat. Hist. nat. Paris* **278**, 1–16 (1975).
36. Halstead, L. B. in *Problems of Phylogenetic Reconstruction* (eds Joysey, K. A. & Friday, A. E.) 159–196 (Academic, London, 1982).
37. Sansom, I. J., Smith, M. P., Armstrong, H. A. & Smith, M. M. *Science* **256**, 1308–1311 (1992).
38. Stensiö, E. A. *Nobel Symposium*, 13–71 (Almqvist & Wiksell, Stockholm, 1968).
39. Novitskaya, L. *Trudy Palaeont. Inst. Akad. Nauk SSSR* **196**, 1–178 (1983).
40. Janvier, P. & Bileck, A. *Zool. Scripta* **8**, 287–296 (1979).
41. Jarvik, E. *Basic Structure and Evolution of the Vertebrates* (Academic, London, 1980).
42. Moy-Thomas, J. A. & Miles, R. S. *Palaeozoic Fishes* 3rd edn (Chapman & Hall, London, 1971).
43. Bunker, S. & Machin, K. E. in *Biomechanics in Evolution* (eds Rayner, J. M. V. & Wootton, R. J.) 113–129 (Cambridge Univ. Press, Cambridge, 1991).
44. Jarvik, E. *Annals Soc. r. Zool. Belg.* **94**, 11–95 (1965).
45. Young, G. C. *Palaeontographica* **167**, 10–76 (1980).
46. Pan, J. & Wang, S.-T. in *Early Vertebrates and Related problems of Evolutionary Biology* (eds Chang, M. M., Liu, Y. H. & Zhang, G. R.) 299–333 (Science Press, Beijing 1991).

ACKNOWLEDGEMENTS. We thank P. Ahlberg (Oxford University) for reading this manuscript and the numerous discoverers and researchers for making this review desirable.

LETTERS TO NATURE

The cosmological deceleration parameter estimated from the angular-size/redshift relation for compact radio sources

K. I. Kellermann

National Radio Astronomy Observatory, 520 Edgement Road, Charlottesville, Virginia 22903-2475, USA

IN cosmological models based on the standard Friedmann–Robertson–Walker geometry, the apparent flux density or angular size of standard candles or standard rods varies with redshift in a way that depends on the deceleration parameter q_0 . (Open universes have $q_0 < 0.5$; closed universes have $q_0 > 0.5$.) At low redshift, however, observational errors are much greater than the differences in q_0 expected for different cosmological models, while at high redshift observational uncertainties, particularly at optical wavelengths, and apparent systematic evolutionary changes in sources obscure the expected geometrical effects. Here I show that measurements by very-long-baseline interferometry (VLBI) of compact radio sources associated with active galaxies and quasars may be largely free of evolutionary effects even at substantial redshifts. The relation between angular size and redshift for a sample of these sources indicates a value of q_0 close to 0.5, corresponding to cosmological density near the critical value.

Hoyle pointed out in 1958 that in conventional Friedmann cosmologies in a non-empty universe ($q_0 > 0$) the angular size, θ , of a 'metric rod' has a minimum near a redshift ($z = 1$) and increases at higher redshifts¹. Observational attempts to use the θ – z relation to distinguish between cosmological models have, however, been hampered at optical wavelengths by the difficulty of precisely measuring angular diameters at high redshifts, and

at radio wavelengths by the apparent change in linear dimensions with redshift and luminosity.

At optical wavelengths, angular-size measurements are difficult because the characteristic size of galaxies at $z \approx 1$ is only a few seconds of arc, comparable with the 'seeing' disk, and it is difficult to define a true metric diameter^{2–4}. Hoyle¹, using Cygnus A as a model ($\theta \approx 100$ arcsec, $z = 0.057$), suggested that the angular separation of the lobes of double radio sources might provide a good 'standard rod'. But at $z \approx 1$, an angular resolution of at least a few arcseconds is needed to measure the lobe separation of highly redshifted double radio sources.

By the 1970s, radio interferometers had achieved sufficient resolution to examine the θ – z relation for radio galaxies and quasars. Miley⁵ and Legg⁶ demonstrated the continuity of measured angular dimensions for radio galaxies and quasars, thus providing convincing evidence for the cosmological origin of quasar redshifts. But, curiously, the observed θ – z relation appeared to follow a simple $1/z$ law, apparently inconsistent with any simple conventional cosmology. Later work^{7,8} provided stronger evidence for the $1/z$ dependence. At redshifts near 2, the angular separations obtained for the lobes of double radio sources were about an order of magnitude smaller than expected from extrapolation of the low-redshift data in standard cosmological models. To save standard cosmologies, it is necessary to assume either that the component separation decreases with cosmic epoch, or that there is an inverse relation between linear size and radio power, so that in a flux-limited sample the more luminous (smaller) sources are preferentially selected at high redshift.

By using large complete samples of radio galaxies and quasars, one can separate the variation of linear size with redshift at a constant luminosity, P , from the variation of linear size with luminosity at a constant redshift. The observed θ – z relation may then be interpreted within standard Friedmann cosmologies as a change of linear, size l , with redshift and luminosity given by

$$l \approx 400(1+z)^{-3}(P/10^{27})^{1/3} \text{ kpc}$$

with P the luminosity in W Hz^{-1} and $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (refs 9–11).

The apparent systematic change in angular size with redshift and luminosity can be reproduced by models based on the propagation of jets through the host galaxy, in an evolving halo, intercluster or intergalactic medium^{12,13}. But it would be a coincidence if the systematic changes in component separation with redshift just cancelled the expected geometric effects of standard models, and this has raised questions about whether the redshift of galaxies and quasars is really due to the expansion of the Universe³.

Compact radio sources observed with VLBI may be free from these systematic evolutionary effects. Their characteristic lifetimes are only a few tens or hundreds of years, much less than the age of the Universe even at the early cosmic epochs corresponding to the highest measured redshifts. Moreover, compact radio sources are smaller than any host galaxy so the radio structure should be independent of the intergalactic or intercluster medium, which is systematically different at high z . Finally, unlike the extended radio sources, compact sources are not randomly oriented in the sky; instead, because of relativistic beaming, they are thought to be aligned close to the line of sight within a narrow range of projection angles centered on an angle $\theta \approx 1/2\gamma$ where $\gamma = (1 - v^2/c^2)^{-1/2}$ (ref. 14). Because such sources are aligned close to the line of sight, small changes in θ correspond to relatively large changes in projected size.

Compact radio sources are also attractive for tests of cosmological models because they are usually identified with quasars, which in general have high redshifts, so the differences between models may be more easily distinguished than for the extended double-lobe sources, which are mostly lower-redshift radio galaxies. As a consequence, however, there are few compact sources at low redshift to define the $1/z$ part of the θ - z diagram. Moreover, both the fraction of sources identified with quasars and the average radio luminosity increase with redshift, which may introduce systematic errors in the analysis of the dependence of angular size on redshift.

Probably the most serious problem with using compact sources is that unlike the extended double-lobe radio sources, their size is not unambiguously defined. In general the most distant jet components are the weakest and most diffuse, and have the

steepest spectra¹⁵, so the apparent angular size of a compact radio source measured with VLBI will depend on sensitivity, angular resolution and frequency. Because observations made at a frequency ν_0 refer to an emitted frequency $(1+z)\nu_0$, where the jets are smaller, an effect of this type (K-correction) may cause high-redshift sources to appear systematically smaller than low-redshift sources.

I have investigated the dependence of angular size on redshift using a sample of 82 compact radio sources taken from the literature which meet the criteria discussed below. About three-quarters of these are taken from five VLBI surveys^{16–20}. The angular extent of each source is defined as the distance between the core and the most distant component whose peak brightness exceeds 2% of the core brightness. These dimensions were determined from published contour diagrams based on VLBI observations at 6 cm wavelength made with a resolution of about 1.5 milliarcsec and a dynamic range of at least 100:1.

In principle, it would be desirable to use an image for each source made at a frequency $\nu_0(1+z)$ to ensure that there is not a systematic change in apparent size with redshift due to the $(1+z)$ shift in frequency. With an array of fixed antennas, however, the change in resolution and sensitivity to extended features with lower surface brightness would vary with redshift, introducing even larger systematic errors. By defining the angular size as the separation of bright components, the effect of observing at a fixed frequency should be small, and is in the sense of an apparent reduction in the measured size of the high-redshift sources, or a decrease in the apparent value of q_0 .

In an attempt to confine the sample to a homogeneous group with similar properties, I have limited the analysis to sources with a luminosity greater than $10^{24} \text{ W Hz}^{-1}$. This is roughly the luminosity that separates radio-loud from radio-quiet quasars²¹ and is also the luminosity at which there is a change in morphological structure²² as well as an observed break in the radio luminosity function of radio galaxies. This restriction to high-luminosity sources does, however, limit the number of low-redshift objects in the sample. I excluded from the sample known BL Lac objects which may be systematically smaller than quasars (because they are thought to be nearly oriented along the line of sight), compact central components of extended double-lobed quasars, and peaked spectrum sources.

The observed θ - z diagram is shown in Fig. 1. The data are divided into seven redshift bins. There were three unresolved sources included in the sample, and for these an angular size of 1 mas was assumed. For another six sources, the resolution was too poor to identify the secondary component clearly, and the component separation was estimated from the size of the image. The unresolved and slightly resolved sources do not seem to be concentrated near any particular redshift interval, so the results are insensitive to the way in which these data are treated. Moreover, a similar analysis using median sizes, which are insensitive to assumptions of the size of the unresolved and slightly resolved sources, gives similar results.

As expected from Friedmann cosmological models with $q_0 > 0$, the observed mean angular size is essentially independent of redshift in the range $0.5 < z < 3$. At a redshift of three, the data lie above the extrapolated $1/z$ value by a full order of magnitude. This is perhaps the best direct observational support for the predictions of Friedmann cosmologies uncontaminated by the evolution of the sample objects, and is also direct evidence that the redshift of galaxies and quasars is really due to the expansion of the Universe. The observed relation between angular size and redshift is consistent with an Einstein-de Sitter Universe with a deceleration parameter $q_0 = 0.5$.

These results provide direct evidence on the largest observable scales that the mass density of the Universe does not greatly differ from the critical value, $\Omega = 1$, separating open and closed universes. They are consistent with the predictions of inflationary cosmologies. $\square$

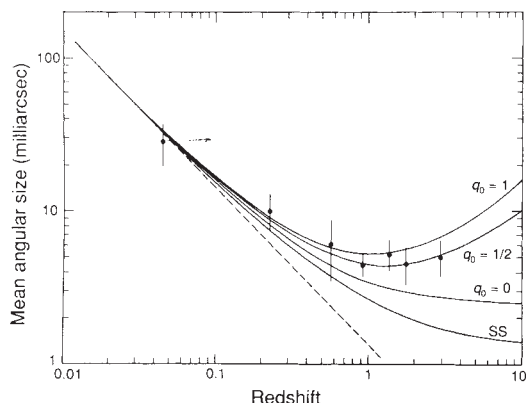


FIG. 1 Mean angular size against redshift for 82 compact sources. From left to right, individual points are based on 8, 4, 11, 13, 20, 13 and 10 sources, respectively. The error bars represent the spread in angular size for each redshift bin. The solid curves represent the expected dependence for Friedmann cosmological models for a standard source with a component separation of 41 parsec and deceleration parameters q_0 of 0, 1/2 and 1 and the steady-state model (SS); the dashed line shows $1/z$ law observed for the separation of double-lobed extended sources.

Received 8 September; accepted 13 November 1992.

1. Hoyle, F. in *Paris Symp. Radio Astronomy*, IAU Symp. No 9 (ed. Bracewell, R.) 529–532 (1959).
2. Sandage, A. R. *Astrophys. J.* **133**, 355–392 (1961).
3. Sandage, A. R. *Astrophys. J.* **173**, 485–499 (1973).
4. Sandage, A. R. *A. Rev. Astr. Astrophys.* **26**, 561–630 (1988).
5. Miley, G. K. *Mon. Not. R. astr. Soc.* **152**, 477–489 (1970).
6. Legg, T. H. *Nature* **226**, 65–67 (1970).
7. Wardle, J. F. C. & Miley, G. K. *Astr. Astrophys.* **30**, 305–315 (1974).
8. Kapahi, V. K. in *Observational Cosmology*, IAU Symp. No. 124 (ed. Hewitt, A., Burbidge, G. & Fang, L. Z.) 251–265 (1987).
9. Oort, M. J. A., Katgert, P., Steeman, F. W. M. & Windhorst, R. A. *Astr. Astrophys.* **179**, 41–59 (1987).
10. Singal, A. K. *Mon. Not. R. astr. Soc.* **233**, 87–113 (1988).
11. Kapahi, V. K. *Astr. J.* **97**, 1–9 (1989).
12. Gopal-Krishna & Wiita, P. J. *Astrophys. J.* **373**, 325–335.
13. Subramanian, K. & Swarup, G. *Mon. Not. R. astr. Soc.* **247**, 237–243.
14. Cohen, M. H. in *BL Lac Objects* (ed. Maraschi, L., Maccacaro, T. & Ulrich, M.-H.) 13–21 (1989).
15. Pearson, T. J. in *Parsec Scale Radio Jets* (ed. Zensus, A.) 1–12 (1990).
16. Pearson, T. J. & Readhead, A. C. S. *Astrophys. J.* **328**, 114–142 (1988).
17. Wilkinson, P. N., Polatidis, A., Readhead, A. C. S., Xu, W. & Pearson, T. J. in *Sub Arc Second Radio Astronomy* (ed. Davis, R. J.) (Cambridge Univ. Press, in the press).
18. Gurvits, L. I. et al. *Astr. Astrophys.* **260**, 82–88 (1992).
19. Ekhardt, A. et al. *Astr. Astrophys. Suppl. Soc.* **67**, 121–146 (1987).
20. Wehrelle, A. et al. *Astrophys. J.* **391**, 589–607 (1992).
21. Kellermann, K. I., Sramek, R., Schmidt, M. & Shaffer, D. B. *Astr. J.* **98**, 1195–1207 (1989).
22. Fanaroff, B. L. & Riley, J. M. *Mon. Not. R. astr. Soc.* **167**, 31p–35p (1974).

ACKNOWLEDGEMENTS. The NRAO is operated by Associated Universities, Inc., under cooperative agreement with the NSF.

Pulsed X-rays from the Vela pulsar

H. Ögelman^{*†}, J. P. Finley^{*} & H. U. Zimmermann[†]

^{*} Department of Physics, University of Wisconsin-Madison, 1150 University Avenue, Madison, Wisconsin 53706, USA

[†] Max-Planck-Institut für Extraterrestrische Physik, D-8046 Garching, Germany

THE Vela pulsar, PSR0833-45, is a young (10^4 yr) nearby neutron star emitting pulsed radiation in radio, optical and γ -ray bands^{1–5}. Soft X-ray imaging by the Einstein⁶ and Exosat⁷ observatories has revealed a point-like source, at the position of the pulsar, embedded in a compact nebula ~ 2 arcmin in diameter, but the search for pulsed X-ray emission has proved difficult. There were claims in 1973 of a positive detection^{8,9} but subsequent observations have failed to confirm them^{6,10}. Here we report an unambiguous detection, by means of the Rosat satellite¹¹, of pulsed X-ray emission from the Vela pulsar, and thus put this long-standing enigma to rest. The pulsed signal is soft, appearing mainly at energies < 1 keV. The Rosat observations resolve the two sources of emission, and show that the point-like emission centred on the pulsar is soft, whereas the emission from the compact nebula is hard. Despite the common prejudice that it is the archetypal young pulsar embedded in a supernova remnant, these observations show that Vela more closely resembles older (10^5 yr) pulsars.

With their strong magnetic fields and short periods, pulsars are expected to generate high voltages and accelerate high-energy electrons and positrons in their magnetospheres which in turn should radiate in X- and γ -rays. Furthermore, the hot surface of a cooling neutron star should be a source of soft X-rays. Up to the present, only a handful of pulsars have been observed to emit energetic photons. It is clearly of interest to observe these objects in the X-ray band and study their spectral and timing characteristics. The Vela pulsar is particularly interesting in that it is one of the closest 'active' neutron stars which, considering its 10^4 -yr age and 89-ms period, should emit X-rays both from magnetospheric processes and initial cooling.

Rosat observed the Vela pulsar during two sets of pointings about 8 months apart, 1991 April 22.0–23.9 and 1991 December 4.7–19.8, with effective exposure times of 13,900 s and 18,500 s respectively. For the timing analysis, we selected all the photons within a radius of $42''$ centred on the pulsar yielding 48,167 and 61,742 photons with consistent count rates of 3.4 s^{-1} . Recognized software errors in the timing were corrected in the analysis. The

photon arrival times of the selected events were corrected to the Solar System barycentre in the barycentric dynamical time unit TDB using the JPL DE200 ephemeris¹² and the radio position J2000 $\alpha = 08 \text{ h } 35 \text{ min } 20.680 \text{ s}$, $\delta = -45^\circ 10' 35.70''$ (J. H. Taylor, personal communication). The accuracy of the Rosat spacecraft clock is consistent to within ~ 0.5 ms over time intervals of a few weeks, although over longer timescales there could be systematic deviations as large as 4 ms. Owing to this long-term clock drift, we period-folded the photon times using the PSR0833–45 radio ephemeris (J. H. Taylor, personal communication) for the April and December data sets separately. Table 1 summarizes the radio ephemeris parameters used. Throughout the observations, the arrival times of the radio pulses were known to an accuracy better than one millisecond. Consequently, there was no need for any period search; each X-ray photon could be directly phase-related to the radio. The resulting phase plots are shown in Fig. 1a and b, where phase 0 is the radio phase accurate to ± 0.05 .

We evaluated the significance of the pulsations with the Z_n^2 test¹³. The main advantage of this test is that it is independent of binning, and the increase of Z_n^2 as a function of n , the number of summed harmonics, shows the harmonic content of the pulse shape. For both April and December observations, the signal increased up to 3 harmonics to Z_3^2 values of 31.2 and 50.1 respectively. Considering that for uniformly distributed phases the values of Z_3^2 should be distributed like χ^2 with 6 degrees of freedom, the probability of random occurrence for the April observation is 2.3×10^{-5} and for December 4.5×10^{-9} . The signal disappears as we change the period of folding or the direction of the accepted events from those expected. We conclude that in both data sets there is strong evidence for pulsations. On the other hand, the two pulse shapes appear different and out of phase. We checked the consistency of the two data sets being drawn from the same parent distribution with a χ^2 test as we varied the phase delay between the two. The results indicate that the two sets are compatible (at the 90% confidence level) if the December data had an additional positive offset of 4–13 ms. Considering that the long-term absolute Rosat time determination could easily have an error of ± 2 ms, we think it likely that we are seeing the effect of timing errors rather than real changes in pulse shape or phase. The fraction of the counts

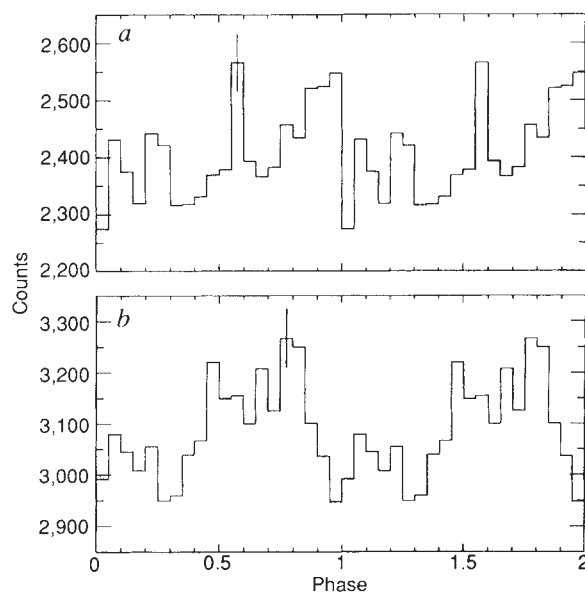


FIG. 1 Folded light curves of the Vela pulsar photons in the energy range 0.06 to 2.4 keV for (a) the April 1991 data set, (b) the December 1991 data set. Phase 0 and 1 refer to the radio pulse with an estimated accuracy of ± 0.05 .

TABLE 1 Radio pulsar data used in pulsation analysis

Date	Epoch* (TDB JD)	ν_0 (Hz)	$\dot{\nu}$ (Hz s ⁻¹)
1991 April 22.0–23.9	2448369.821485628	11.198852134351	$-1.557662 \times 10^{-11}$
1991 December 4.7–19.8	2448604.913344683	11.198564513900	$-1.565618 \times 10^{-11}$

* The epoch is also the radio arrival time at the Solar System barycentre.

that are pulsed in Fig. 1b is $4.4 \pm 1.1\%$, which is consistent with the upper limit of $\sim 9\%$ for broad pulse shapes established by the Einstein observatory⁶. The general features of the soft X-ray light curve of the Vela pulsar include a smaller pulse in the phase range 0.0–0.3 and a larger double-peaked pulse in the phase range 0.4–1.0. The X-ray pulse is complex and does not match the shapes of the γ -ray, optical or radio pulse. Compared with the radio phase, the bulk of the X-ray pulse resides in the phase region 0.5–1.0 whereas the optical and the γ -ray pulse dominate the region 0.1–0.6 (ref. 4). We should point out that this is the first attempt at absolute phase comparisons with the Rosat data. There is still concern that the calibration of the spacecraft clock may contain an offset (P. Predehl, personal communication).

To investigate the spectral characteristics of the pulses, the point source and the compact nebula, we first had to determine the contributions from each component as a function of radius and energy. To this end, we constructed a maximum likelihood procedure in which we fitted the observed radial count density profiles out to 3' in 24 energy bins (covering channels 6–240) with the following parameters: (1) the counts $N_{ps}(E)$ from the point source distributed with the point response function of Rosat at that energy; (2) the counts $N_{di}(E)$ from the compact nebula distributed with various radial nebular distribution functions determined iteratively; (3) a flat background count density $BG(E)$. The results of the fits were satisfactory and yielded the raw count distribution as a function of Rosat energy channel for the point-like and diffuse components separately (Fig. 2). The point-source spectrum is softer than the nebular component. The scale size of the nebula is compatible with the high-angular-resolution Einstein High-Resolution-Imager data⁶ and the Exosat Channel Multiplier Array data⁷. The soft point-source spectrum is supported by the earlier Exosat filter spectroscopy results⁷. Using these energy-dependent functions, we estimate that the pulsed fraction of the point-source counts is 11%. With this data set we also did spectral fits to the compact nebula and the point source. The compact nebula is well fitted by either a power-law or thermal bremsstrahlung spectrum; we get unacceptable fits to a black-body spectrum. For the point source, we fitted the data with black-body and power-law models, which give acceptable results. We also fitted this data below 1.2 keV to eliminate the high energy tail. These results are summarized in Table 2. Finally, we did spectral fits to the pulsed counts in five energy channels with sufficient statistical significance. As in the raw spectrum of the point source, there were no pulsed counts above 1.2 keV, again confirming that the pulsed photons were soft. The spectrum best fitting the pulsed photons was a black-body model with parameters similar to the point source. These results are included in Table 3. Again the pulsed bolometric luminosity comes out to be $\sim 10\%$ of the bolometric luminosity of the point source. It is also reassuring that the spectral parameters of the pulsed photons, chosen on the basis of time variation, are almost identical to the point-source photons chosen on the basis of radial count-density fits.

The period (~ 0.089 s) and period derivative of the Vela pulsar yield a characteristic age $\tau \approx 11,000$ yr and a rotational energy loss rate $\dot{E} \approx 7 \times 10^{36}$ erg s⁻¹ (ref. 14). Among the five previously known pulsars that emit pulsed X-rays, the age of the Vela pulsar lies between the young, $\sim 10^3$ -yr ages of the Crab pulsar, PSR0540–69 and PSR1509–58 and the old, $\sim 10^5$ -yr ages of PSR0656+14 and Geminga¹⁵. The young pulsars all have a power-law-type spectrum which is interpreted as the magneto-

spheric emission of the relativistic electrons and positrons^{16,17}. The older pulsars have a soft, black-body-like spectrum with some residual excess at higher energies^{18,19}. Our results show that in X-rays, the Vela pulsar is similar to the older pulsars although it is a factor of 10 younger. The X-ray pulsed fraction of Vela (11%) is also comparable to those of PSR0656+14 (14%; ref. 18) and Geminga (15%; ref. 19). The combination of parameters that determine the transition of the X-ray spectrum from the hard magnetospheric emission to the soft black-body-like emission must have already exceeded the boundary for the Vela pulsar. Black-body fits to the spectrum of the Vela point source give surface temperatures in the range 1.5 – 1.6×10^6 K, which is compatible with some cooling models and reheating processes^{20–22}, but the observed luminosities of 4 – 5×10^{32} erg s⁻¹ imply that the radius of the Vela pulsar is only 3–4 km rather than the expected 10–18 km (at infinity). This discrepancy can be resolved when we take into account the spectral modification of the neutron star atmosphere, including the effect of strong magnetic fields^{18,23–25}. The lower luminosity and similar temperature of the black-body fits to the pulsed flux suggest a radius around three times less for the origin of this component. The high harmonic content of the pulse shape would require a complex surface temperature structure or angle-dependent opacities due to the magnetic field geometry.

The power-law fit to the compact nebula is consistent with the Einstein IPC results⁶, although the IPC fit includes both the compact nebula and the point source together. The factors determining the size of this distinct region are still not clear⁷. Its luminosity must derive from the wind of the pulsar, but the fact that the luminosity of the compact nebula is only $\sim 3 \times 10^{-4}$ of the spin-down luminosity implies that this region is

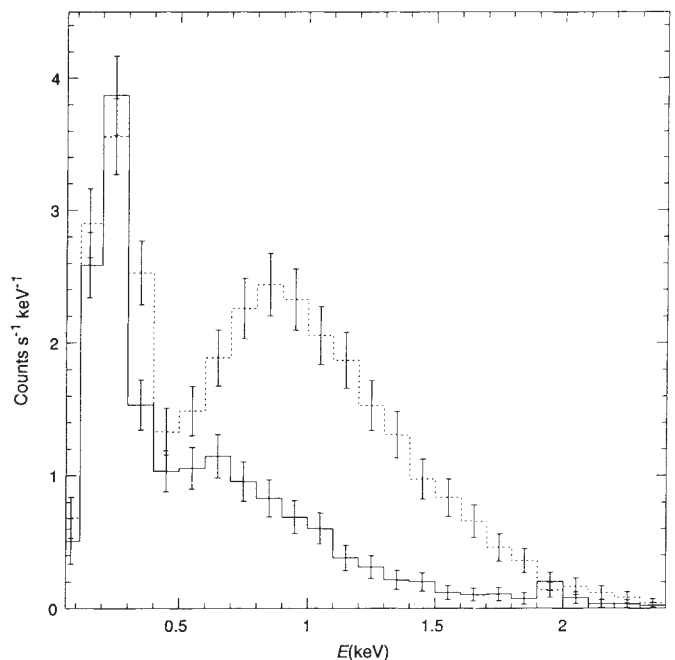


FIG. 2 Energy distribution of the point source (solid lines) and the compact nebula photons (dashed lines) as derived from the energy-dependent radial-count-density distributions. The errors are estimated from the range of acceptable radial-count-density models of the compact nebula.

TABLE 2 Spectral models and fit parameters

Model	N_{H} (10^{20} cm^{-2})	Model parameter	L_{x}^* ($10^{32} \text{ erg s}^{-1}$)
Compact nebula (0.06–2.4 keV)			
Thermal brems.	2.0 ± 0.5	$kT = 1.6 \pm 0.2 \text{ keV}$	14
Power-law	3.0 ± 0.5	photon index = -2.0 ± 0.2	18
Point source (0.06–2.4 keV)			
Black-body	0.5 ± 0.4	$kT = 0.15 \pm 0.01 \text{ keV}$	4.3
Power-law	4.3 ± 1.0	photon index = -3.3 ± 0.4	26
Point source (0.06–1.2 keV)			
Black-body	0.9 ± 0.6	$kT = 0.13 \pm 0.02 \text{ keV}$	4.8
Power-law	4.4 ± 1.5	photon index = -3.3 ± 0.8	26
Pulsed photons (0.06–2.4 keV)			
Black body	1.0 ± 0.8	$kT = 0.14 \pm 0.02 \text{ keV}$	0.5

* The luminosity at 500 pc in the energy band 0.1–2.4 keV. For black-body fits it is the bolometric luminosity.

remarkably transparent to the low-frequency dipole waves and the relativistic hydromagnetic wind generated at the magnetosphere of the pulsar.

This detection of the Vela pulsar, together with the recent Rosat discoveries of Geminga¹⁵ and PSR0656+14 (ref. 18) as pulsed soft X-ray sources, implies that Rosat can study the X-ray spectrum and pulse profiles of 10^4 – 10^5 -yr-old neutron stars within a few kiloparsecs from us. Other such discoveries together with observations at other wavelengths may allow us to construct

a consistent picture of magnetospheric emission and the cooling properties of neutron stars. □

Received 21 October; accepted 23 November 1992.

1. Large, M. I., Vaughan, A. E. & Mills, B. Y. *Nature* **220**, 340–341 (1968).
2. Wallace, P. T. *et al. Nature* **266**, 692–694 (1977).
3. Thompson, D. J., Fichtel, C. E., Kniffen, D. A. & Ögelman, H. B. *Astrophys. J.* **214**, L17–L18 (1977).
4. Kanbach, G. *et al. Astr. Astrophys.* **90**, 163–169 (1980).
5. Tümer, O. T. *et al. Nature* **310**, 214–216 (1984).
6. Harnden, F. R. Jr, Grant, P. D. & Seward, F. D. *Astrophys. J.* **299**, 828–838 (1985).
7. Ögelman, H. & Zimmermann, H.-U. *Astr. Astrophys.* **214**, 179–185 (1989).
8. Harnden, F. R. Jr, Johnson, W. N. III & Haynes, R. C. *Astrophys. J.* **172**, L91–L94 (1972).
9. Harnden, F. R. Jr & Gorenstein, P. *Nature* **241**, 107–108 (1973).
10. Knight, F. K., Matteson, J. L., Peterson, L. E. & Rothschild, R. E. *Astrophys. J.* **260**, 553–560 (1982).
11. Trümper, J. *et al. Adv. Space Res.* **2**, 241–249 (1983).
12. Standish, M. *Astr. Astrophys.* **114**, 297–302 (1982).
13. Buccheri, R. *et al. Astr. Astrophys.* **128**, 245–251 (1983).
14. Manchester, R. N. & Taylor, J. H. *Astr. J.* **86**, 1953–1973 (1981).
15. Ögelman, H. in *Proc. Workshop Physics of Isolated Pulsars* (Cambridge Univ. Press, Cambridge, in the press).
16. Cheng, K. S., Ho, C. & Ruderman, M. *Astrophys. J.* **300**, 522–539 (1986).
17. Harding, A. K. *Astrophys. J.* **245**, 267–273 (1981).
18. Finley, J. P., Ögelman, H. & Kızıloğlu, Ü. *Astrophys. J.* **394**, L21–L24 (1992).
19. Halpern, J. P. & Holt, S. S. *Nature* **357**, 222–224 (1992).
20. Nomoto, K. & Tsuruta, S. *Astrophys. J.* **312**, 711–726 (1987).
21. Shibasaki, N. & Lamb, F. K. *Astrophys. J.* **346**, 808–822 (1989).
22. Page, D. & Applegate, J. H. *Astrophys. J.* **394**, L17–L20 (1992).
23. Romani, R. W. *Astrophys. J.* **313**, 718–726 (1987).
24. Miller, M. C. in *Proc. Workshop Physics of Isolated Pulsars* (Cambridge Univ. Press, Cambridge, in the press).
25. Ventura, J., Shibanov, Yu. A., Zavlin, V. E. & Pavlov, G. G. in *Proc. Workshop Physics of Isolated Pulsars* (Cambridge Univ. Press, Cambridge, 1992).

ACKNOWLEDGEMENTS. We thank Günther Hasinger for the time-corrected Rosat data. This work was supported by NASA LTSA Research Program.

Photoinduced electron transport across a lipid bilayer mediated by C₇₀

Kuo Chu Hwang & David Mauzerall

Rockefeller University, 1230 York Avenue, New York, New York 10021, USA

ELECTRON transport across a membrane is central to photosynthesis, to mitochondrial respiration and to the design of molecular systems for solar energy conversion. Relatively few synthetic molecules, however, have been shown to facilitate transport of electrons across a lipid bilayer^{1–3}. We report here that C₇₀ can act as both a photosensitizer for electron transfer from a donor molecule and a mediator for electron transport across a lipid bilayer membrane. The steady-state photocurrent density obtained from the C₇₀-bilayer system is about 40 times higher, at comparable light intensities, than that of the carotene-porphyrin-quinone system², previously the most efficient artificial system. The C₇₀-bilayer system has a quantum yield of about 0.04, while the stability (tens of minutes) and turnover number (electrons transported per C₇₀ before decay) of 10^3 are one to three orders of magnitude greater than those of other systems^{1–3}. We anticipate that other higher fullerenes may also provide the basis for efficient transmembrane electron-transport systems.

C₆₀ has been shown^{4–6} to be a good electron acceptor when photoexcited. Cast films of C₆₀ on metal electrode surfaces have been found to act as an *n*-type semiconductor, but the dark current is at least an order of magnitude larger than the photocurrent⁷. Polyvinyl carbazole doped with C₆₀ is an efficient photoconductor, in which holes in the carbazole valence band act as the charge carriers and the C₆₀ molecules act as the photoactive electron acceptors⁸. We have reported elsewhere⁹ that both C₆₀ and C₇₀ embedded within a lipid membrane can act as efficient electron acceptors from photoexcited amphoteric zinc porphyrin absorbed to the membrane; the fullerene anions carry the charges deeply into the bilayer, enhancing the photovoltage. We

show here that these fullerenes can transpose electrons completely across the membrane.

Figure 1a shows a transmembrane photocurrent of 4 nA produced by continuous illumination of a lipid bilayer containing C₇₀ which separates two aqueous compartments with ascorbate and sodium anthraquinone 2-sulphonate (hereafter, ascorbate [C₇₀] AQS, where | | represents the two membrane-water interfaces). The sign of the photovoltage⁹ and of the photocurrent together with redox potentials (see below) show that electrons are transported from aqueous ascorbate to photoexcited C₇₀ in the membrane, then to AQS in the other aqueous interface. The redox potentials are as follows: $E^0_{\text{ascorbate(ox)}/\text{ascorbate(red)}} = -0.21 \text{ V}_{\text{SCE}}$ (ref. 10), $E^0_{\text{C}_{70}/\text{C}_{70}^-} = -0.43 \text{ V}_{\text{SCE}}$ (ref. 11), $E^0_{\text{C}_{70}/\text{C}_{70}^{2-}} = +1.14 \text{ V}_{\text{SCE}}$ (ref. 5) (the first reduction potential¹¹ and the triplet energy^{12,13} of C₆₀ and C₇₀ are about the same). The presence of molecular oxygen suppresses >95% of the photocurrent. Our photovoltage measurements⁹ indicate that interfacial electron transfer from (60 mM) ascorbate to (0.8 mM) C₇₀ occurs on a timescale of ~80 μs. The lifetime of photoexcited C₇₀ is 0.67 ns (ref. 14) for the singlet state and 51 ms (ref. 12) for the triplet state, so triplet C₇₀ must be the reactive species. No photocurrent (<0.1 pA) was observed whenever C₇₀ was omitted from the bilayer, proving that C₇₀ is the transmembrane electron carrier. On long irradiation (30 min, Fig. 1b), the photocurrent began to decay after ~2 min, and the decay rate increased with increasing light intensity. The photocurrent remains at its final low level after 30 min in the dark. The cause of the slow photocurrent decay is currently unclear. The total charge transported by C₇₀ across the bilayer is 4 μC (or 2.5×10^{13} electrons). We estimate that there are $\leq 4 \times 10^{10}$ molecules of C₇₀ in the bilayer, assuming the ratio of C₇₀ to lipid remains constant from the decane-toluene solution to the bilayer. The turnover number of the system (one electron per C₇₀) is thus ~ 10^3 .

When a membrane is permeable to electrons, a Nernst potential is generated by the redox potential difference of a redox pair in the two aqueous compartments^{1,15}. As shown in Fig. 2, on irradiation the observed potential of the C₇₀-containing membrane is proportional to the calculated Nernst potential difference created by the [ascorbate]_{red}/[ascorbate]_{ox} redox pair, with a slope of 0.72. In contrast, a small, constant (−1 ± 1 mV)

membrane potential was observed in the dark. The results show that the C_{70} -containing membrane is permeable to electrons when illuminated. The magnitude of the slope decreases at lower C_{70} concentration and for weaker incident light intensity. The slope of less than unity may be caused by permeability to other ions^{1,15} or by side reactions of the intermediate ascorbic radicals. Photovoltage measurements⁹ indicate that C_{70} is stable in the bilayer for ≥ 1 s.

To find the limiting factors, we investigated the effects of light, donor, acceptor and C_{70} concentrations on the magnitude of the photocurrent. As shown in Fig. 3a, both light and ascorbate are limiting factors. The photocurrent increases linearly with light intensity from 7 to 220 mW cm^{-2} , and increases monotonically with the ascorbate concentration from 10 to 100 mM. In contrast, the photocurrent is saturated at 0.1 mM AQS (data not shown), so the concentration of acceptor is not limiting. The effect of the C_{70} concentration is shown in Fig. 3b. The photocurrent reaches a plateau at ~ 0.7 mM, probably because of limited C_{70} solubility in the bilayer. Overall, the photocurrent is limited by light and by interfacial electron transfer on the donor side.

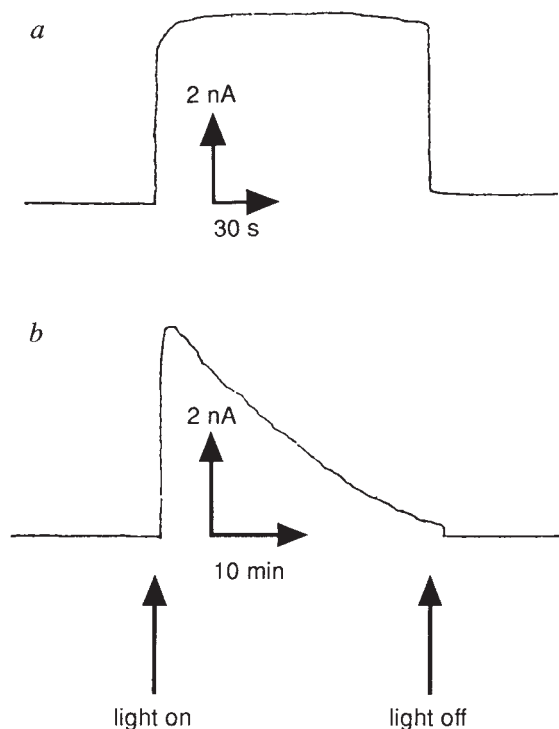


FIG. 1 Positive transmembrane photocurrent obtained from a C_{70} -containing lipid bilayer. C_{70} (Materials and Electrochemical Research Corporation) was chromatographically purified¹⁹, dissolved in toluene and added to a 4% w/w 1,2-diphytanoyl-sn-glycero-3-phosphocholine/decane solution (decane:toluene 2:1). The final C_{70} concentration was 0.4 mM. A planar lipid bilayer was formed in a hole of 1.5 mm diameter in a thin Teflon sheet dividing two aqueous compartments which contained 1 M NaCl, 10 mM HEPES, pH7.0. Oxygen was removed by addition of glucose, glucose oxidase and catalase (ref. 20). The bathing solution was the same in all experiments. Control experiments show that the glucose oxidase can serve as a weak electron acceptor (ref. 21). The transmembrane photocurrent was measured using two calomel electrodes clamped at zero volts by an operational amplifier²², and the output plotted on a recorder. Unless otherwise specified, the incident light was from a 300-W slide projector filtered by a 1-cm, 1 M CuSO_4 aqueous solution, to give 2.4 mW cm^{-2} light at wavelength $\approx 400\text{--}600$ nm on the membrane area of $\sim 1.1 \times 10^{-2} \text{ cm}^2$. The dark current was < 0.1 pA. The system was 80 mM ascorbate [0.4 mM C_{70}] 0.2 mM AQS, and with irradiation times of (a) 112 s and (b) 30 min. Ascorbate was on the negative electrode side and AQS on the positive electrode side. The peak current was 4.3 nA in a and 4.5 nA in b. Typically, a C_{70} -containing membrane has a resistance of $\sim 9 \text{ G}\Omega$ in the dark.

We estimated the quantum yield of transmembrane photocurrent for the C_{70} -membrane system in the following way. A 300-pA photocurrent was observed from a system of 100 mM ascorbate [0.4 mM C_{70}] 0.3 mM AQS, irradiated by light of 8 mW cm^{-2} (through an interference filter, wavelength $\lambda = 540 \pm 6$ nm). Given an optical cross-section of $0.6 \text{ \AA}^2$ for C_{70} at 540 nm (ref. 11), and assuming that the C_{70} to lipid ratio in the decane-toluene solution remains unchanged on forming the bilayer, we calculated a quantum yield of transmembrane electron transport of 0.04. The agreement of the action spectrum (within $\pm 15\%$ (r.m.s.)) with the absorption of C_{70} from ~ 400 to 600 nm (data not shown) shows that C_{70} is the photoactive species.

The C_{70} -bilayer system may conduct by two different electron-transport mechanisms: simple diffusion of the fullerene anions or electron hopping among fullerene molecules, possibly in the form of aggregates. The mismatch between the nearly spherical C_{70} structure and the lipid bilayer chain structure suggests that free diffusion would be slow. A chain structure of C_{70} aggregates across the bilayer could be favoured. There is evidence for an ion-chain mechanism of conductance by large hydrophobic anion and cation pairs across bilayer¹⁶. In addition, electron hopping between C_{70} molecules in aggregates may be faster than usual because of the near-spherical symmetry giving better overlap in electronic wavefunctions. Measurements below the lipid gel transition temperature may distinguish between these two mechanisms¹⁷, as the transmembrane diffusion of fullerene anions will be essentially stopped whereas electron hopping will not. So far such experiments have been unsuccessful.

We have also found that C_{60} is an efficient transmembrane electron carrier, with a quantum yield of ~ 0.1 (data not shown). In general, the observed photocurrent from the ascorbate [C_{60}] AQS system is three- to fivefold smaller than that of the corresponding C_{70} system. The fact that the optical cross-section of C_{60} is about eightfold smaller than that of C_{70} in the visible region outweighs the higher quantum yield. Because of their similarities in structure, electronic levels and redox properties, we predict that other higher fullerenes will be equally good transmembrane electron carriers.

Attempts to mimic photosynthetic energy storage in the form of charge separation with membranes or films are often limited

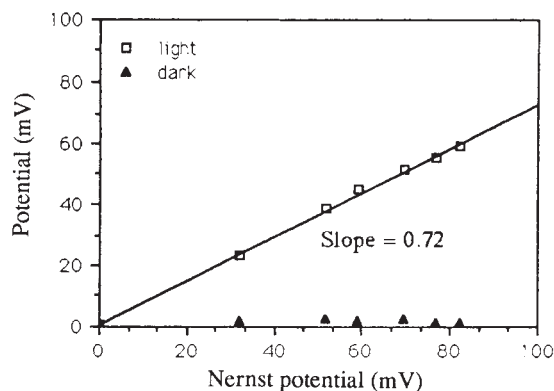


FIG. 2 Steady-state membrane potential ($\square$, in light; $\blacktriangle$, in dark) measured as a function of the calculated Nernst potential difference across the membrane, created by the difference in concentration of ascorbate in the two aqueous compartments. The C_{70} concentration, in the lipid-forming solution is 0.7 mM. The ascorbate concentration was maintained constant at 2.1 mM on the negative aqueous compartment, and the ascorbate concentration was adjusted from 2.1 to 50.1 mM on the positive side. To produce the oxidized form of ascorbate, we added 0.1 mM ferricyanide to both aqueous compartments before adding the ascorbate. A comparable slope was obtained in the absence of ferricyanide and also with 5 mM ascorbate as reference. The incident light was 18 mW cm^{-2} on the membrane area from a 300-W slide projector (unfiltered). The membrane potential (rise time < 1 s, instrument limited) was measured with calomel electrodes and a Keithley high-impedance ($10^{14} \Omega$, model 617) electrometer.

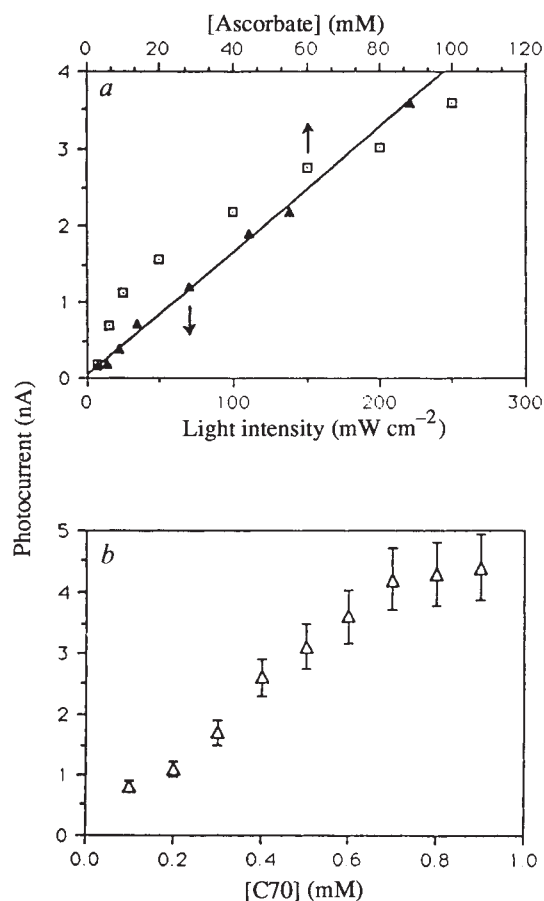


FIG. 3 a, Photocurrent measured as a function of ascorbate concentration (□) and incident light intensity (▲). The system consisted of [ascorbate] 0.7 mM, [C₇₀] 0.3 mM, AQS. Either the ascorbate concentration was varied from 0 to 100 mM and the incident light intensity was constant at 220 mW cm⁻² ($\lambda \approx 400$ –600 nm), or the ascorbate concentration was constant at 100 mM and the light intensity was varied using glass filters. The photocurrent polarity and the locations of ascorbate and AQS are the same as those in Fig. 1. b, Photocurrent measured as a function of C₇₀ concentration with 30 mM ascorbate in one aqueous compartment, 0.2 mM AQS in the other, and incident light intensity of 220 mW cm⁻² ($\lambda \approx 400$ –600 nm).

by the difficulties of transferring charge across the membrane. The use of fullerenes or of graphitic nanotubes²³, may reduce this limitation. The stability of lipid bilayers can be vastly improved through the use of microporous filters as holders¹⁸. The self-organizing feature of these films is a desirable characteristic for molecular electronics. □

19. Ajie, H. *et al.* *J. phys. Chem.* **94**, 8630–8633 (1990).
20. Benesch, R. E. & Benesch, R. *Science* **111**, 447–448 (1953).
21. Schuhannk, W., Ohara, T. J., Schmidt, H. L. & Heller, A. *J. Am. chem. Soc.* **113**, 1394–1397 (1991).
22. Hong, F. T. & Mauzerall, D. *J. electrochem. Soc.* **123**, 1317–1324 (1976).
23. Iijima, S. *Nature* **354**, 56–58 (1991).

ACKNOWLEDGEMENTS. We are grateful for critical discussions with T. Marinetti. This research was supported by the NIH.

CO₂, CH₄ and N₂O flux through a Wyoming snowpack and implications for global budgets

R. A. Sommerfeld*, A. R. Mosier† & R. C. Musselman*

*USDA Forest Service, 240 West Prospect Road, Fort Collins, Colorado 80526, USA

†USDA Agricultural Research Service, PO Box E, Fort Collins, Colorado 80522, USA

INCREASING atmospheric concentrations of the three main greenhouse gases—carbon dioxide, methane, and nitrous oxide—account for about 70% of anticipated global warming¹, but the production–consumption budgets are not balanced for any of these gases². Snow can cover between 44 and 53% of the land area of the Northern Hemisphere³ and may be several metres deep in alpine and sub-alpine regions for more than half the year. Most trace-gas budgets assume that trace-gas exchange stops when soil is snow covered or soil temperatures drop to $\sim 0^\circ\text{C}$ (refs 4, 5). Thus alpine and sub-alpine soils are generally considered to be net sinks for atmospheric CO₂. Some reports^{6,7}, however, suggest that soil microorganisms beneath the snow continue to respire at temperatures close to 0°C . Here we present evidence that the soils under alpine and sub-alpine snowpacks emit CO₂ and N₂O and take up atmospheric CH₄ throughout the snow-covered period. These fluxes represent an important part of the annual trace-gas budget for these ecosystems.

Carbon dioxide is important environmentally and biologically. It is the primary gas involved in the exchange of carbon between the atmosphere and the Earth. As a component of the atmosphere, carbon dioxide is an important greenhouse gas contributing $\sim 45\%$ of all greenhouse forcing^{1,4}. A large amount of carbon is returned to the Earth from the atmosphere by photosynthesis and is subsequently released from biota to the atmosphere by respiration.

The contribution of CO₂ from alpine or arctic regions in winter has not been considered to be important in calculations of global carbon balances. However, high concentrations of carbon dioxide have been reported^{8–10} at the base of snowpacks in arctic and temperate climates. At present, estimates of fluxes of CO₂ from natural ecosystems and from snow are incomplete, making it difficult to evaluate the controlling processes (see for example ref. 11).

Methane (CH₄) and nitrous oxide (N₂O), also important greenhouse gases, are responsible for $\sim 20\%$ of anticipated annual global warming¹. Despite considerable research, large uncertainties exist in the global budget of both CH₄ and N₂O. Both gases have important soil sources, about 70% for CH₄ and 90% for N₂O (ref. 4). The observation of relatively high rates of CH₄ consumption in aerobic soils suggests an important soil sink term in the global CH₄ budget^{5,12–14}. It is generally assumed that snow-covered soils do not contribute either positively or negatively to N₂O and CH₄ atmospheric budgets^{4,5}. If soils below the snow remain biologically active relative to CO₂ and N₂O production and CH₄ consumption, then a portion of the annual budget of these gases may be missing in model calculations. Gas fluxes can be calculated from concentration profiles because the porosity of the snow can be estimated to within about 10% (ref. 15). Here we assume the gas flux is the result

Received 26 May; accepted 7 December 1992.

1. Feldberg, S. W., Armen, G. H., Bell, J. A., Chang, C. K. & Wang, C. B. *Biophys. J.* **34**, 149–163 (1981).
2. Seta, P. *et al.* *Nature* **316**, 651–655 (1985).
3. Seta, P., Bienvenue, E., Maillard, P. & Momenteau, M. *Photochem. Photobiol.* **49**, 537–543 (1989).
4. Sension, R. J., Szarka, A. Z., Smith, G. R. & Hochstrasser, R. M. *Chem. Phys. Lett.* **185**, 179–183 (1991).
5. Arbogast, J. W., Foote, C. S. & Kao, M. *J. Am. chem. Soc.* **114**, 2277–2279 (1992).
6. Kamat, P. V. *J. Am. chem. Soc.* **113**, 9705–9707 (1991).
7. Miller, B. *et al.* *J. Am. chem. Soc.* **113**, 6291–6293 (1991).
8. Wang, Y. *Nature* **356**, 585–587 (1992).
9. Hwang, K. C. & Mauzerall, D. *J. Am. chem. Soc.* **114**, 9705–9706 (1992).
10. Clark, W. M. *Oxidation-Reduction Potentials of Organic Systems*, 470 (Williams & Wilkins, Baltimore, 1960).
11. Allemann, P. M. *et al.* *J. Am. chem. Soc.* **113**, 1050–1051 (1991).
12. Wasielewski, M. R., O'Neill, M. P., Lykke, K. R., Pellin, M. J. & Gruen, P. M. *J. Am. chem. Soc.* **113**, 2774–2776 (1991).
13. Hung, R. R. & Grabowski, J. J. *J. phys. Chem.* **95**, 6073–6075 (1991).
14. Kim, D., Lee, M., Suh, Y. D. & Kim, S. K. *J. Am. chem. Soc.* **114**, 4429–4430 (1992).
15. Iliani, A. & Mauzerall, D. *Biophys. J.* **35**, 79–92 (1981).
16. Drain, M. & Mauzerall, D. *Biophys. J.* **63**, 1556–1563 (1992).
17. Krasne, S., Eisenman, G. & Szabo, G. *Science* **174**, 412–415 (1971).
18. Mountz, J. M. & Tien, H. T. *Photochem. Photobiol.* **28**, 395–400 (1978).

TABLE 1 Average positive flux of gases to the atmosphere

Site	Year	CO ₂ (10 ⁻² mol m ⁻² day ⁻¹)			N ₂ O (10 ⁻⁷ mol m ⁻² day ⁻¹)			CH ₄ (10 ⁻⁶ mol m ⁻² day ⁻¹)		
		Flux	Range	<i>n</i>	Flux	Range	<i>n</i>	Flux	Range	<i>n</i>
Sa	1991	4.6	1.8–12.4	12	8.5	2.4–23.4	8	–6.3	–16.6–2.6	9
	1992	3.6	1.4–3.9	4	5.2	1.9–6.4	3	–2.9	–7.0–1.5	4
Sb	1991	6.2	2.2–13.7	13	8.5	5.2–21.5	8	–15.5	–31.4–5.8	7
	1992	4.2	3.5–5.5	4	5.9	4.4–15.7	3	–2.2	–7.4–0.3	4
Aa	1991	1.2	1.0–3.7	10	5.0	2.8–18.8	7	–1.8	–3.4–0.3	6
	1992	1.7	0.8–3.6	5	3.4	2.4–4.0	3	–2.2	–2.6–0.9	6
Ab	1991	1.7	1.0–3.1	10	2.1	–0.3–5.6	7	–4.4	–9.2–0.9	6
	1992	1.7	0.7–4.3	5	3.1	–2.3–3.8	2	–3.3	–4.8–1.6	5

of simple, gradient-driven diffusion and a single porosity for the snow.

The sampling locations for this study were in the Glacier Lakes Ecosystem Experiments site (GLEES) in the Snowy Range of southeastern Wyoming, west of Laramie at about 41° 20' N, 106° 20' W. Samples were taken from two alpine sites (Aa, Ab) and two sub-alpine sites (Sa, Sb). The alpine sites, separated by about 10 m, are at the ecotone between alpine and sub-alpine, at an elevation of 3,286 m. The location is on an exposed, rocky ridgetop with little soil development, and has sparse alpine vegetation with small, widely scattered trees. The sub-alpine sampling sites, separated by about 3 m, are in a small opening in a sub-alpine spruce-fir forest at an elevation of 3,182 m. This location is a typical sub-alpine meadow. Geology, soils, vegetation and topography details of GLEES are given in ref. 16.

Gas collectors⁷ were installed in the ground at 15 cm below the surface, at the surface, and in the snowpack at 5, 20, 50 and 100 cm above the ground at each site. Samples were drawn using 20-ml nylon syringes fitted with two-way nylon stopcocks. Serum bottles (20 ml), capped with butyl rubber stoppers, were evacuated and filled for transport to the Fort Collins Agricultural Research Service laboratory for N₂O analyses and for CH₄ analyses in 1991. Another 20 ml were drawn and the syringes brought to a field laboratory ~10 km from the sample locations, for analysis of CO₂ and CH₄ (1992) by gas chromatography within five hours of sample collection. The 1991 N₂O and CH₄ samples were stored for periods up to 3 months before analysis by gas chromatography^{17,18}.

Up to seven depths were sampled for each of the four sites between 2 April and 28 May 1991 and between 12 March and 14 May 1992. The sampling intervals were between 1 and 7 days in 1991 and roughly every 14 days in 1992. Figure 1 shows the average concentrations of CO₂, N₂O and CH₄ in the snow and ground during the sample periods, at the four sites.

The average CO₂ profiles are in good agreement with the preliminary results of ref. 7. In contrast to those measurements, the initial samples showed high CO₂ concentrations at all levels indicating that respiration was occurring before appreciable melt water flowed from the snow. We attribute the difference to poor sampling technique in the initial samples used in ref. 7. The CH₄ concentrations for 1991 are less accurate than the 1992 data because of storage problems.

On 7 May 1991, a snow pit was dug near the sub-alpine location to obtain an estimate of the snow porosity. The snow depth at the pit was 205 cm, about 50 cm deeper than at the two sub-alpine sample sites. However, the stratigraphy did not show any large changes in porosity. Discontinuous ice lenses were located at 143, 158 and 184 cm. The thickness-weighted average porosity of the upper 100 cm of snow was ~0.58. For most of the profiles, gradients in the layer above 100 cm are similar to or smaller than those from 50 to 100 cm (Fig. 1). This indicates that the ice lenses did not have a significant effect on the flux. Furthermore, it is our experience that ice lenses disappear early in the melt process after the snow contains a significant amount of liquid water. Snow pits in 1992 indicated that the upper snow layers had similar porosities.

Because biological activity was suspected in the snow, the pit dug on 7 May 1991 was sampled for biological material. The profile of the combined counts of fungus, algae and bacteria showed a strong peak at 50 cm, corresponding to the inflection point in the CO₂ and N₂O profiles. Fungus filaments and spores formed the main fraction of the biota sampled.

Table 1 gives calculated fluxes of the gases at each site. They were calculated assuming simple diffusion, average gradients from 100 cm to the surface, average snow depths estimated from limited measurements and an average porosity in the top layer from the snow pit. Because the snow pit was dug later in the season after the snow had compacted, the calculated fluxes are probably conservative.

$$J_g = D_g(d[g]/dz)f$$

where J_g is the gas flux, $d[g]/dz$ is the vertical concentration gradient and f is the porosity. The diffusion coefficient D_g was 0.139 cm² s⁻¹ for CO₂ and N₂O and 0.22 for CH₄, an experimental value for CO₂ and estimated values for the other gases¹⁹. Molar volumes of the gases were corrected for temperature and pressure. Replications and calibrations indicate that the concentration data are accurate to ±3% for CO₂ and for CH₄ in 1992. This is the precision of the measurements of N₂O and of CH₄ in 1991 which, because of longer storage, are thought to have an accuracy of about ±10%. The ranges of the data and the numbers of samples are also shown in Table 1.

The soil temperatures at 20 cm below the soil surface in a forest clearing about 0.5 km from the alpine location dropped below 0 °C in mid-November and slowly decreased to –0.2 °C. At the sub-alpine location, soil temperatures at 20 cm dropped to 1.5 °C in mid-November and decreased to 0.5 °C before the mid-June snowmelt when the soil temperature increased rapidly at both sites. Although soil temperature remained within ~1 °C of freezing during the sampling period, soil microbial activity did not cease. The soil continuously served both as a source of CO₂ and N₂O, which evolved to the atmosphere through the snow, and as a sink for atmospheric CH₄. As the main biological activity was in the soil and the soil temperature did not change much during the entire winter, we believe the results from the sampling periods are representative of the entire winter.

The amount of carbon mobilized as CO₂ at the sub-alpine sites ranges from 2.9 to 14.5 mol m⁻² for a snow-covered period of 235 days. Litter fall in a lodgepole pine forest near our sites was estimated to contribute between 4.45 and 7.43 mol m⁻² yr⁻¹ (ref. 20). Litter fall should be considerably lower in our non-forested sites. The total amount of carbon fixed in temperate forests in the United States is estimated to be 13.3 mol m⁻² yr⁻¹ (ref. 21, 22). Less carbon would be fixed at high elevation locations such as those we sampled, as low temperature, short growing season and low light levels under the snow cover limit production. Although these comparisons are approximate, they show that the amount of carbon respired in the winter in snow-covered environments can be an important fraction of the total carbon budget.

The profiles also indicate that respiration occurred in the snow at about the 50-cm level in 1991. Respiration in the snow may be related to the high fungus and bacterial counts at this

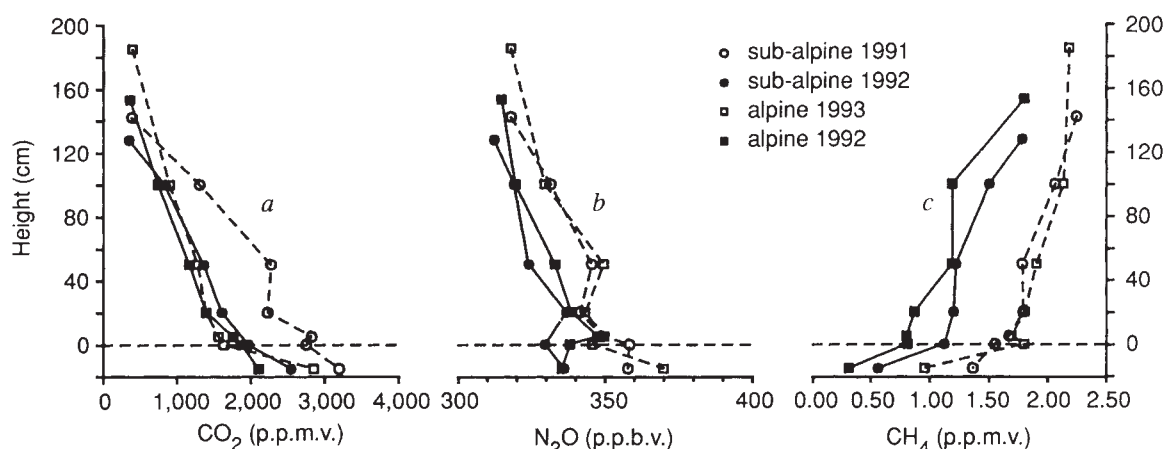


FIG. 1 Average concentrations of CO_2 (a), N_2O (b) and CH_4 (c) in the snow at four sites for the periods sampled in 1991 and 1992.

level. CO_2 production in the snow significantly increased the flux at sub-alpine sites in 1991 but not in 1992 or at the alpine sites.

During the summer of 1991, between mid-June (when the snow cover melted) and mid-October (when the snow cover returned), N_2O and CH_4 gas fluxes were sampled over a meadow located ~50 m from the sub-alpine location. N_2O emissions, measured at 1- to 7-day intervals, averaged $1.34 \times 10^{-7} \text{ mol m}^{-2} \text{ d}^{-1}$ during the 130-day collection period (A.R.M., L. K. Klemmedtion, R.C.M. and R.A.S., unpublished data). The N_2O emissions through the snow averaged $7.0 \times 10^{-7} \text{ mol m}^{-2} \text{ d}^{-1}$ at the sub-alpine location and $3.4 \times 10^{-7} \text{ mol m}^{-2} \text{ d}^{-1}$ at the alpine location over a 235-day period.

We estimate CH_4 flux from the atmosphere to the soil under the snow to be 6.7×10^{-6} and $2.9 \times 10^{-6} \text{ mol m}^{-2} \text{ d}^{-1}$ at the sub-alpine and alpine locations respectively. In the adjacent meadow, CH_4 flux into the soil averaged $7.4 \times 10^{-6} \text{ mol m}^{-2} \text{ d}^{-1}$ throughout the snow-free period (A.R.M., L. K. Klemmedtion, R.C.M. and R.A.S., unpublished data). These data suggest that the soil serves as a larger CH_4 sink while under the snowpack (235 days) than during the snow-free time (130 days). In the Colorado short grass steppe during the winter, when no snow cover existed and soil temperature was $< -2^\circ\text{C}$, the soils continued to take up CH_4 from the atmosphere at the rate of $6.8\text{--}13.6 \times 10^{-5} \text{ mol m}^{-2} \text{ d}^{-1}$ (ref. 13).

Our data show that the soil microflora are active even when they are snow-covered and near 0°C . N_2O is produced as a result of nitrification and denitrification in soil processes that proceed at higher rates at higher temperature, but have been observed at -2°C in supercooled soil²³. Emissions from the soil through snow continued throughout the period of snow cover. Methane uptake under the snow seems to represent a considerable portion of the CH_4 flux during the year. Soil respiration under the snow continued through the sampling period and represents oxidation of more than 25% of the estimated carbon fixed in the ecosystem during the growing season. Soils having surface horizon temperatures near the freezing point represent a considerable fraction of the land surface in the Northern Hemisphere during the winter³. The fact that similar concentrations of CO_2 have been measured at the bases of snowpacks in Utah¹⁰ and Fairbanks and Barrow, Alaska^{8,9}, indicates that similar fluxes occur in these widely spaced, dissimilar regions. Observations^{6,8-11,14} indicate that the data presented here are representative of such systems and that the winter fluxes of CO_2 , N_2O , and CH_4 are important in the global budgets of these gases. □

- Barry, R. G. in *Encyclopedia of Earth System Science* vol. 1, 517-524 (Academic, San Diego, 1992).
- Bouwman, A. F. in *Soils and the Greenhouse Effect* (Wiley, Chichester, 1990).
- Stedler, P. A., Bowden, R. D., Melillo, J. M. & Aber, J. D. *Nature* **341**, 314-316 (1989).
- Taylor, B. R. & Parkinson, D. *Soil Biol. Biochem.* **20**, 657-665 (1988).
- Sommerfeld, R. A., Musselman, R. C., Reuss, J. O. & Mosier, A. R. *Geophys. Res. Lett.* **18**, 1225-1228 (1991).
- Kelley, J. J., Weaver, D. F. & Smith, B. P., *Ecology* **49**, 358-361 (1968).
- Coyne, P. I. & Kelley, J. J. *Geophys. Res.* **79**, 799-802 (1974).
- Solomon, D. K. & Cerling, T. E. *Wat. Resour. Res.* **23**, 2257-2265 (1987).
- Kling, G. W., Kipphut, G. W. & Miller, M. C. *Science* **251**, 198-201 (1991).
- Keller, M., Mirtre, M. E. & Stallard, R. F. *Glob. biogeochem. Cycles* **4**, 21-27 (1990).
- Mosier, A., Schimel, D., Valentine, D., Bronson, K. & Parton, W. *Nature* **350**, 330-332 (1991).
- Whalen, S. C. & Reeburgh, W. S. *Nature* **346**, 160-162 (1990).
- Martinez, J. & Rango, A. *J. Hydrol.* **84**, 197-219 (1986).
- Musselman, R. C. *Rocky Mt. Forest & Range Expt. Station Gen. Tech. Rep.* (in the press)
- Bronson, A. Mosier, A. R. *Biol. Fert. Soils* **11**, 116-120 (1991).
- Mosier, A. R. & Mack, L. *Soil Sci. Soc. Am. J.* **44**, 1121-1123 (1980).
- Perry, R. H., Chilton, C. H. & Kirkpatrick, S. D. in *Chemical Engineers' Handbook* (McGraw-Hill, New York, 1963).
- Fahey, T. J. *Ecol. Monogr.* **53**, 51-72 (1983).
- Marland, G. in *Global Climate Change Linkages: Acid Rain, Air Quality, and Stratospheric Ozone* (Elsevier, New York 1989).
- Kinsman, J. D. & Marland, G. *82nd A. Mtg Air and Waste Management Assoc.* Pap. 89-148.5 (Pittsburgh, 1989).
- Dorland, S. & Beauchamp, E. G. *Can. J. Soil Sci.* **71**, 293-303 (1991).

Numerical simulation of self-organized stone stripes

B. T. Werner* & B. Hallet†

* Center for Coastal Studies 0209, Scripps Institution of Oceanography, La Jolla, California 92093-0209, USA

† Quaternary Research Center, AK-60, University of Washington, Seattle, Washington 98195, USA

GEOMETRICALLY regular stripes of stones are found on many unvegetated alpine and polar hillslopes; known as 'sorted stripes' because of the characteristic textural sorting between surface stones and fine-grained soil, they contrast markedly with the lack of order typical of natural landscapes. The spacing of the stripes can range from centimetres to metres (about 10-20 times the average stone diameter^{1,2}), with individual stripes extending down-slope for many tens of metres (Fig. 1). A variety of formative mechanisms have been proposed³, but it is still unclear how such orderly stripes can arise spontaneously, and what dictates the spacing. Here we present two-dimensional computer simulations in which the displacement of surface stones is governed by the preferential growth of needle-ice in stone-free regions of soil during each freezing cycle. Regular patterns of stones and soil that closely resemble natural stripes are found to emerge spontaneously from the model as a result of identifiable feedbacks between soil texture, ice growth and local slope.

The recurrent frost-induced displacement of surface stones is a primary characteristic of hillslopes mantled by sorted stripes.

Received 24 June; accepted 8 December 1992.

- Rohde, H. *Science* **248**, 1217-1219 (1990).
- IPCC in *Climate Change: The IPCC Scientific Assessment*, (eds Houghton J. T., Jenkins C. J. & Ephraums J. J.) Introduction, 36-37. (Cambridge Univ. Press, Cambridge, (1990).

Field data indicate that the growth and decay at the ground surface of fibrous ice, known as needle ice, are very effective in displacing stones by raising them from the surface during the freezing phase and dropping them on thawing⁴⁻⁶. Upward frost-induced stone displacements can convey subsurface stones to the ground surface; sorted stripes have, however, been observed on slopes where subsurface displacement was negligible⁶, bringing into question models for the formation of sorted stripes that focus on lateral subsurface rearrangements of stones due either to uneven ice growth or uneven thawing. The former may reflect for example, incipient variations in soil texture⁸, the latter may be dictated, for example, by free convection of water through the thawed soil⁹. Other mechanisms for the formation of sorted stripes, such as those involving preferential stone accumulation in pre-existing water-eroded rills⁴ or in desiccation cracks that tend to be aligned downslope¹⁰, have not received widespread support because they lack supportive evidence from diverse field sites.

Here we present the initial results of a numerical model that simulates stochastic displacements of stones by needle-ice activity on a slope with a bimodal mixture of stones and fine-grained soil. Needle ice develops under conditions in which the moist surface soil is brought to the freezing temperature by radiative cooling¹¹. The tendency for moisture to remain preferentially in fine-grained soil (on unvegetated hillslopes) and the need for capillarity to import water for needle-ice growth suggest that more needle ice should develop on the surface in soil domains than in stony domains; this hypothesis is supported by field evidence⁵. Moreover, needle-ice toppling is expected to

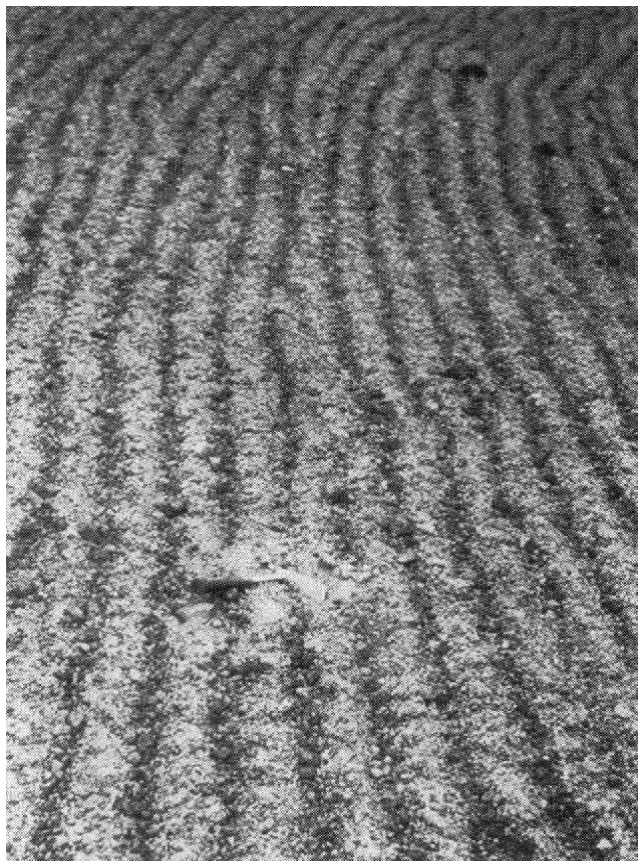


FIG. 1 Sorted stripes spaced 12–15 cm apart extend downslope for tens of metres on a 15° slope near the summit of Mauna Kea, Hawaii; stone bands consist of centimetre-sized pebbles. Stripes of this type occur in many parts of the world on unvegetated hillslopes mantled by a bimodal mixture of stones and soil, and subject to frequent freeze/thaw events³. View is upslope. Photo courtesy of S. Porter.

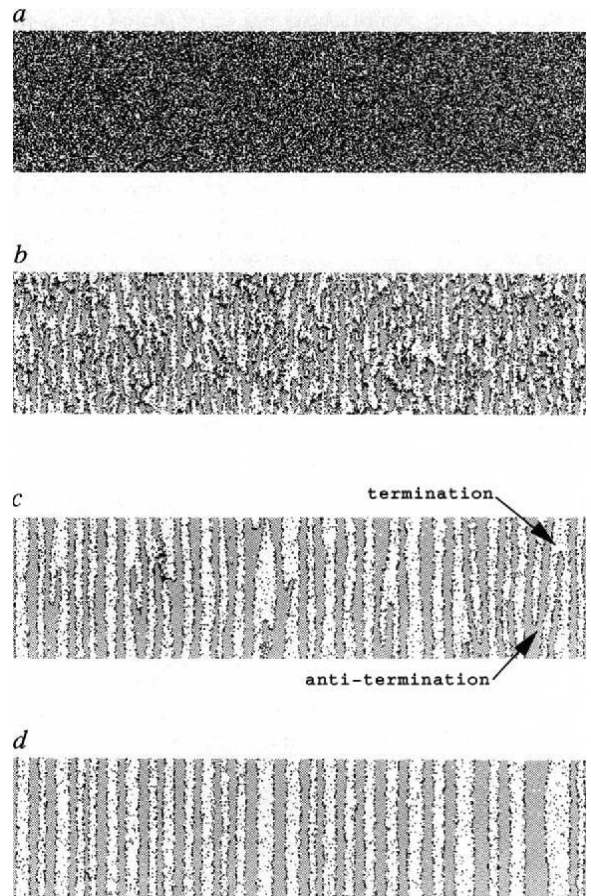
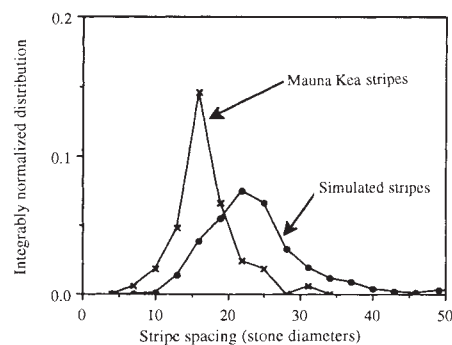


FIG. 2 Plan view of sorted-stripe computer simulation with the downslope direction oriented downwards in the figures. A two-dimensional square grid (640 × 160 cells here) with periodic boundary conditions forms the surface in the simulations. A cell can be occupied by one stone (denoted by a square which is black if it has moved in the last needle-ice cycle, varying to grey if it has not moved for 40 or more needle-ice cycles) or can represent exposed soil (denoted by a white square). Stones are arranged initially on the grid in random positions. The system is evolved forward in time by moving individual stones to neighbouring cells according to stochastic rules based on the likelihood of needle-ice formation and the computed local slope. The probability of transport to one of the up to eight empty neighbouring cells depends on the gravitational potential energy change ΔE associated with that move; the most negative ΔE is favoured. ΔE is a sum of the potential energy change associated with the mean slope, and that associated with the local slope caused by the preferential elevation of soil domains, taken to be proportional to the stone concentration gradient. Stone concentration is averaged over a square region 3 × 3 stone diameters. Lacking a mechanical model for needle-ice growth and collapse on a rough surface, the probability of transport to a neighbouring cell is calculated according to the Boltzmann distribution ($\sim e^{-\Delta E/kT}$), which facilitates introduction of an effective temperature (kT) specifying the magnitude of random fluctuations in the direction of stone transport. Mean and local slopes appear in the model relative to kT only. For example, in the simulation displayed, if the mean slope is 10°, then random variations in the surface slope are of the order of 1° and the maximum magnitude of the local slope (measured relative to the mean hillslope) is 2°. Each cell is polled once for stone uplift and toppling in a simulated needle-ice cycle. *a*, Initial random stone configuration with stones filling 50% of the cells. *b*, After 400 simulated needle ice cycles. *c*, After 4,000 cycles. *d*, After 14,000 cycles. Because of the limitation to two dimensions, stones within a stone band rarely move, unlike stones in natural stone bands. Therefore, simulated needle-ice cycles overestimate the effective amount of time elapsed by underestimating stone mobility. Assuming that the rate of sorted-stripe formation scales with stone mobility and using the observation that stones in the soil band move downslope twice as fast as stones in the stone band for natural sorted stripes⁶, a rough estimate of the effective number of needle-ice cycles is the number of simulated needle-ice cycles times the ratio of simulated to actual observed downslope transport rates. Under these assumptions, *b*, *c* and *d* roughly correspond to 50, 300 and 1,100 natural needle-ice cycles respectively.

FIG. 3 Stripe spacing distribution for sorted stripes on Mauna Kea (55 samples) and for simulated sorted stripes shown in Fig. 2 combined with data from the same simulation with different initial conditions (for a total of 54 independent samples of spacing). For both, stripe spacing was measured from one soil-to-stone transition to the next along a cross-slope transect; in the simulations, the spacing distribution was averaged over transects from top to bottom of the simulation. Stripe spacing for the Mauna Kea data was normalized using the diameter of stones that occupy the greatest area in the stone bands, 0.85 cm (actually, the centre of a plateau ranging from 0.5 to 1.2 cm). Mean and standard deviation for Mauna Kea stripe spacing are 16.9 and 4.2 stone diameters, and for the simulations, 23.9 and 5.5, respectively. We expect simulated stripes to be larger in spacing because of the absence of sub-stone diameter transport distances and the limitation to eight discrete directions of needle-ice topple. The ratio of standard deviation to mean spacing is similar in both cases: 0.23 on Mauna Kea and 0.25 in the simulations. This ratio for soil and stone band widths taken separately is larger for both: 0.32 for stone bands and 0.38 for soil bands on Mauna Kea; 0.31 for stone bands and 0.40 for soil bands in the simulations.



transport stones further in soil domains than in stone domains⁶, where neighbouring stones inhibit stone transport. As ice needles are expected to grow perpendicular to the local surface (the local freezing plane), the most likely topple direction is along the local fall line. The local slope is a combination of the mean inclination of the hillslope, the slope created on the margins of soil domains as they are preferentially raised by subsurface ice growth¹², and random factors such as surface roughness or arrangement of adjacent particles.

We investigate the transport mechanisms underlying the development of sorted stripes using two-dimensional, plan-view numerical simulations of stochastic transport of individual stones. In contrast with Ahnert's model¹³ simulating generic particle rearrangements, the present simulations were designed specifically to approximate the physical processes described above. In the simulations, sorted stripes emerge and develop spontaneously (Fig. 2). Random spatial fluctuations in the initial texture evolve into regions of high and low stone concentration slightly larger than the area over which stone concentration is averaged to obtain the local slope. This initial division into stone and soil domains arises through positive feedback. Fluctuations in stone concentration are unstable because of the decreasing mobility of stones with increasing stone concentration, owing both to the interference by neighbouring stones and to limited needle ice growth in stone domains. Relative uplift of soil domains by preferential subsurface ice growth focuses stone transport along the margins of incipient stone domains. Net downslope stone transport leads to a downslope elongation of incipient textural domains, resulting in a series of alternating stone and soil bands (Fig. 2b). The stone bands have finite length, with ends that face upslope (terminations) and downslope (anti-terminations). Increase in stripe spacing is caused by the relative upward (downward) propagation of terminations (anti-terminations), Fig. 2c. Where terminations and anti-terminations meet, they annihilate, eventually producing a well organized stripe pattern (Fig. 2d). This pattern is fairly stable because negative feedback due to preferred downslope transport tends to straighten out bends in stone bands.

Stripe formation requires the mean hillslope and the local slope due to preferential soil domain uplift to be within an order of magnitude of each other and larger than random fluctuations in the slope. Sorted stripe development is robust to variations of all other simulation parameters. With zero mean slope, sorted circles, nets and islands form. In the absence of soil domain uplift, transverse, upslope-propagating stone bands develop. The stripes and other sorted patterns in the simulations are self-organized because they arise from local feedback between stone concentration and stone transport; they are not guided by an externally supplied template.

Two dominant processes dictating the stripe spacing have emerged from the simulations. (1) Stripe spacing increases

through progressive elimination of stripe termination pairs resulting from mismatches in the initial stripe pattern (Fig. 2), until all pairs vanish. This mechanism implies that stripe spacing depends on the minimum domain size needed for significant soil domain uplift. (2) Stripe spacing increases indefinitely at a decreasing rate with time through the creation of termination pairs (initially gaps in a stone band) followed by the propagation and eventual elimination of the terminations from each pair as they annihilate with other terminations. The result is loss of stone stripes and an increase in the mean stripe spacing. Both processes are sensitive to random fluctuations in the stripe edge position, which in the simulations are dominated by the finite stone size, and secondarily by the degree of randomness of stone displacements (the effective temperature, Fig. 2). Simulated stripe spacing scales with stone diameter and is fairly insensitive to stone transport length (roughly the length of needle-ice crystals) unless the transport length exceeds the averaging domain size determining local slope, whereupon the stripe pattern is lost. Future simulations and field observations should help in determining the relative importance of these mechanisms.

In addition to the visible similarity between patterns generated and natural sorted stripes, several properties are shared by actual and simulated stripes, including (1) distribution of stripe spacings with a standard deviation about 1/4 of the mean (Fig. 3); (2) a strong tendency for stones initially in fine-grained soil domains to end up in stone domains, and the very weak tendency for the reverse situation; (3) stochastic stone movements with no correlation between stones; and (4) more rapid downslope stone displacement in the soil than in the stone domains. Extension of the algorithm to three dimensions should allow more quantitative comparisons with field data.

The prime merit of this type of grain-scale simulation is that it provides an effective tool for probing the fundamental physics underlying the development of a class of self-organized landform patterns, of which sorted stripes is one example. The simulations have shown the importance of both the mean hillslope and the local slope due to frost-induced uplift of soil in determining the direction and magnitude of stone transport. The interplay between these two components of the slope is the key to producing self-organized stripes. Simulations of sorted stripes and other landscape features, including beach cusps (B.T.W. and T. M. Fink, manuscript in preparation), wind ripples^{14,15} and dunes, will help improve our understanding of how local feedbacks can give rise to the patterns that frequently adorn the landscape. □

Received 22 September; accepted 26 November 1992.

1. Goldthwait, R. *Quat. Res.* **6**, 27–35 (1976).
2. Noguchi, Y., Tabuchi, H. & Hasegawa, H. *Geogr. Ann.* **A69**, 329–342 (1987).
3. Washburn, A. L. *Geocryology*, 406 (Wiley, New York, 1980).
4. Brockie, W. J. *Proc. 4th New Zealand Geogr. Conf.*, 91–104 (1965).

5. Brookie, W. J. *Proc. 5th New Zealand Geogr. Conf.*, 191–201 (1967).
6. Mackay, J. R. & Mathews, W. H. *Arc. Alpine Res.* **6**, 347–359 (1974).
7. Anderson, S. *Geol. Soc. Am. Bull.* **100**, 609–621 (1988).
8. Hallet, B. *Can. J. Phys.* **68**, 842–852 (1990).
9. Krantz, W. B. *Earth Sci. Rev.* **29**, 117–130 (1990).
10. Gradwell, M. W. *New Zealand J. Sci. Technol.* **38**, 793–806 (1957).
11. Outcalt, S. I. *Water Resources Res.* **7**, 394–400 (1971).
12. Coutard, J. P., Van Vliet-Lanoë, B., Auzet, C. & Auzet, A. V. Z. *Geomorph. N.F.* **71**, 13–23 (1988).
13. Ahnert, F. *Trans. Jap. geomorph. Un.* **2**, 301–312 (1981).
14. Anderson, R. S. *Earth Sci. Rev.* **29**, 77–96 (1990).
15. Forrest, S. B. & Haff, P. K. *Science* **255**, 1240–1243 (1992).

ACKNOWLEDGEMENTS. This work was supported by the U.S. Army Research Office (to University of Washington), the Office of Naval Research, Coastal Sciences (to Scripps Institution of Oceanography) and the NSF, Surficial Processes (to Duke University).

Oblique rifting in a slow-spreading ridge

Olivier Dauteuil & Jean-Pierre Brun

Laboratoire de Tectonique, Géosciences Rennes, Campus de Beaulieu, Avenue du Général Leclerc, 35042 Rennes Cedex, France

IN oceanic rifts, the spreading direction is usually nearly perpendicular to the ridge axis, but at the Reykjanes¹ and Mohns^{2–4} ridges this direction is highly oblique (30° to 40° to the axis). These ridges possess axial valleys containing oblique raised features (highs), corresponding to second- or third-order ridge axis discontinuities⁵. At the Reykjanes ridge, Searle and Laughton¹ observed small highs which they interpreted as volcanic domes. Here we present a fault map of the Mohns ridge obtained from image processing of Seabeam data. We argue that the oblique highs in this ridge are of tectonic origin. The axial valley is interpreted as the surface trace of a deformable band of oceanic lithosphere inherited from an earlier stage of perpendicular spreading. We describe laboratory experiments on brittle–ductile models which support this interpretation, in which stretching at 30° to a deformable band reproduces the *en échelon* horst-and-graben pattern of the Mohns ridge.

The Mohns ridge in the Norwegian Sea (Fig. 1a) trends WSW–ENE and extends for 400 km from the Jan Mayen fracture zone to Knipovitch ridge. The ridge deepens to the northeast as it nears the Iceland hotspot^{7–9}. The Norwegian sea started to open 53 Myr ago with a NNW–SSE relative displacement between the Greenland and European plates. The full spreading rate was 2.5 cm yr^{−1}, and the Mohns ridge formed perpendicular to the spreading direction^{7–10}. A reorganization of North Atlantic plate boundaries 27 Myr ago induced WNW–ESE spreading. From 12 to 5 Myr (ref. 4), the spreading rate slowly decreased to 1 cm yr^{−1}, and then increased to its present value of 1.8 cm yr^{−1}. In spite of these kinematic changes, the ridge retained its old trend and is now oblique to the 110–120° spreading direction (Fig. 1a).

Seabeam data acquired by IFREMER (Département des Géosciences Marines) in 1988 provide 100% coverage of the axial valley and rift shoulders for a 250-km-long segment, and also include 170-km-long profiles perpendicular to the ridge. The axial valley (Fig. 1b) is continuous, but deviates slightly and in places is not clearly visible. Its width ranges from 8 to 15 km. Four oblique topographic highs, located within the rift valley, are spaced at 20 to 45 km intervals, and trend at ~30° to the ridge. These highs separate deep (−3,300 m or deeper) sigmoidal basins whose width is 15 km and whose lengths range from 15 km to 24 km. They have flat floors which most probably are covered by volcanic flows, the seismic data providing no indication of sediments¹¹. Small undulations could correspond to older topographic breaks covered by lava flows. From northeast to southwest along the valley, the two most prominent highs have steep flanks with a central plateau. The third is very sharp, and the fourth rounded and deeper. The structure of the Mohns ridge is different from that of the Reykjanes ridge, where highs

are more closely spaced (~5 km) and are not separated by deep basins. The rift valley shoulders of the Mohns ridge are strongly asymmetric. The northwest shoulder is the shallowest and is made up of slopes with variable trends. They are parallel to the valley trend in the northeast part, trend east–west in the central part and are NNE in the southwest part. The southeast shoulder has a smoother topography with slopes parallel to the axial valley organized into two steps, especially in the southwest part. Similar asymmetry has been observed on a larger scale^{7,8}. The wall disappears at 72.20° N, 2° E, and beyond this point the bathymetry is smoother and the valley evolves progressively to the southeast shoulder without any sharp relief.

We produced a detailed structural map (Fig. 1c) using several bathymetry images obtained by numerical processing of Seabeam data (shaded relief, slope). From this we determined the polarity of fault scarps, and identified landslides and volcanic structures. Within the axial valley, faults are linear or smoothly curved and have a mean 030° trend oblique to the rift valley. Approaching the rift valley walls, faults are strongly curved and become parallel to valley walls. Topographic highs in the axial valley are bounded by large vertical offset faults indicating that they are horsts. The boundary fissures acted as conduits for the volcanism. Whereas the deep basins show little sign of faulting, the steep northwest wall contains many imbricated faults. These are mainly parallel to the axial valley; they are sigmoidal, curved or linear in form, and are often *en échelon* or interconnected. Along the southeast wall, faults are less numerous and have smaller throws. They are mainly left-lateral *en échelon* faults with a mean 030° trend. On the northwest shoulder, faults are generally curved and have a mean 040° trend; on the southeast shoulder, they are more linear with an *en échelon* pattern. At 72.20° N, 2° E, the characteristic 030° fault pattern is interrupted by a band with smooth topography oriented at 110° and bounded by several 110°–130° faults. On both sides of this zone, horsts and basins are displaced in the same direction and with similar magnitude. Three zones of this type can be identified in the survey area (Fig. 1c): we interpret these as transfer zones¹² which accommodated variations in the extent of stretching along the valley.

The distribution of fault directions is asymmetrical and ranges from 000° to 120° with a well defined peak at 035°, a large population between 040° and 100° and a small peak at 110°–120°. When compared to data from experiments on oblique rifting⁶, such a distribution leads us to predict a 105° stretching direction, which is close to the direction of spreading obtained from bulk kinematic reconstructions^{7–9}. Kinematic analysis of fault patterns therefore permits an independent estimate of spreading direction, and is especially useful in regions lacking transform faults.

The above data and our analysis of the faulting suggest that the axial valley is the surface response to a deformable band in the oceanic lithosphere (Fig. 2) at the junction between the Greenland and European plates. The direction of the band was inherited from the earlier stage of rifting, from 53 to 27 Myr ago (anomaly 7), when spreading was perpendicular to the ridge (Fig. 1a). Later spreading at an angle of 30° to the rift produced the oblique rifting pattern of Mohns ridge. Deformation was largely confined to the axial valley whose great width was controlled by ductile deformation at depth (Fig. 2). The deformable band was not wide enough to allow the development of a rift-transform configuration, and produced only an *en échelon* pattern of horsts and grabens.

The above interpretation was tested in small-scale experiments on brittle–ductile models. In earlier experiments⁶, the width-to-thickness ratio of the upper brittle layer of the deformable band was not properly scaled to model a slow-spreading ridge. In the present work, the thickness of the brittle layer was estimated using the graben width (W_g), following the approach of Allemand and Brun¹³. Taking a value of 30° for the frictional angle in brittle crust¹⁴, we calculated that the dip of normal

faults was 60° and the brittle thickness $t_b = W_g/2 \tan 60^\circ$. In the Mohns ridge, where the mean W_g in the valley is 15 km, the brittle layer would be ~ 8 km thick. The width of the deformable band (W_{db}), estimated from the width of the axial valley, is 8–15 km, implying that the width-to-thickness ratio (W_{db}/t_b) of the brittle layer within the deformable band was between 1 and 2. In the experiments, W_{db} was controlled by a band of silicone putty with newtonian rheology at the base of the model (Fig. 3a). The brittle layer above the silicone band was pure quartz sand characterized by a 30° frictional angle. Displacement was applied to the base of the silicone band by two diverging plastic sheets. The divergence obliquity α was controlled by the transverse (v_t) and longitudinal (v_l) velocities applied at the boun-

daries. To comply with the configuration of the Mohns ridge, we adopted a value of 30° for α (consistent with that found by kinematic analysis of faulting) and a direction of v_l corresponding to left-lateral strike slip.

Figure 3b and c shows surface views of two experiments with different amounts of extension. Normal faults, which developed at 15° to the zone of deformation, are organized in left-lateral *en échelon* systems defining an axial valley with an irregular width. With increasing extension, the fault traces progressively change from rectilinear to sigmoidal, and many faults connect to form rhomboidal blocks. The entire pattern (Fig. 3c) is comparable to that observed on Mohns ridge (Fig. 1c). Discontinuities in the fault pattern define linear bands nearly parallel

FIG. 1 a, Schematic map of Norwegian–Greenland Sea. b, Artificially shaded bathymetry of Mohns ridge: light inclined at 45° comes from the northeast. This processing is sensitive to small bathymetric breaks and permits a better interpretation of the bathymetry. It allows fractures to be distinguished from volcanic lineaments, or the erosion effects on topographic features to be estimated. c, Structural sketch map drawn from different numerical images (shaded reliefs, slope). In the key, 1 indicates normal faults with their dip direction; 2, faults without identifiable direction of dip; 3, probable faults; 4, volcanoes; 5, transfer zones. d, Distribution of faults in the Mohns ridge. The histogram is centred on the strike of the rift valley for comparison to the data of Tron and Brun⁶. The lower diagram, in which the angle of obliquity is plotted against fault distribution, comes from experimental data⁶. Thus we determine that the stretching direction (105°) is close to the spreading direction 110° obtained from the kinematic reconstruction.

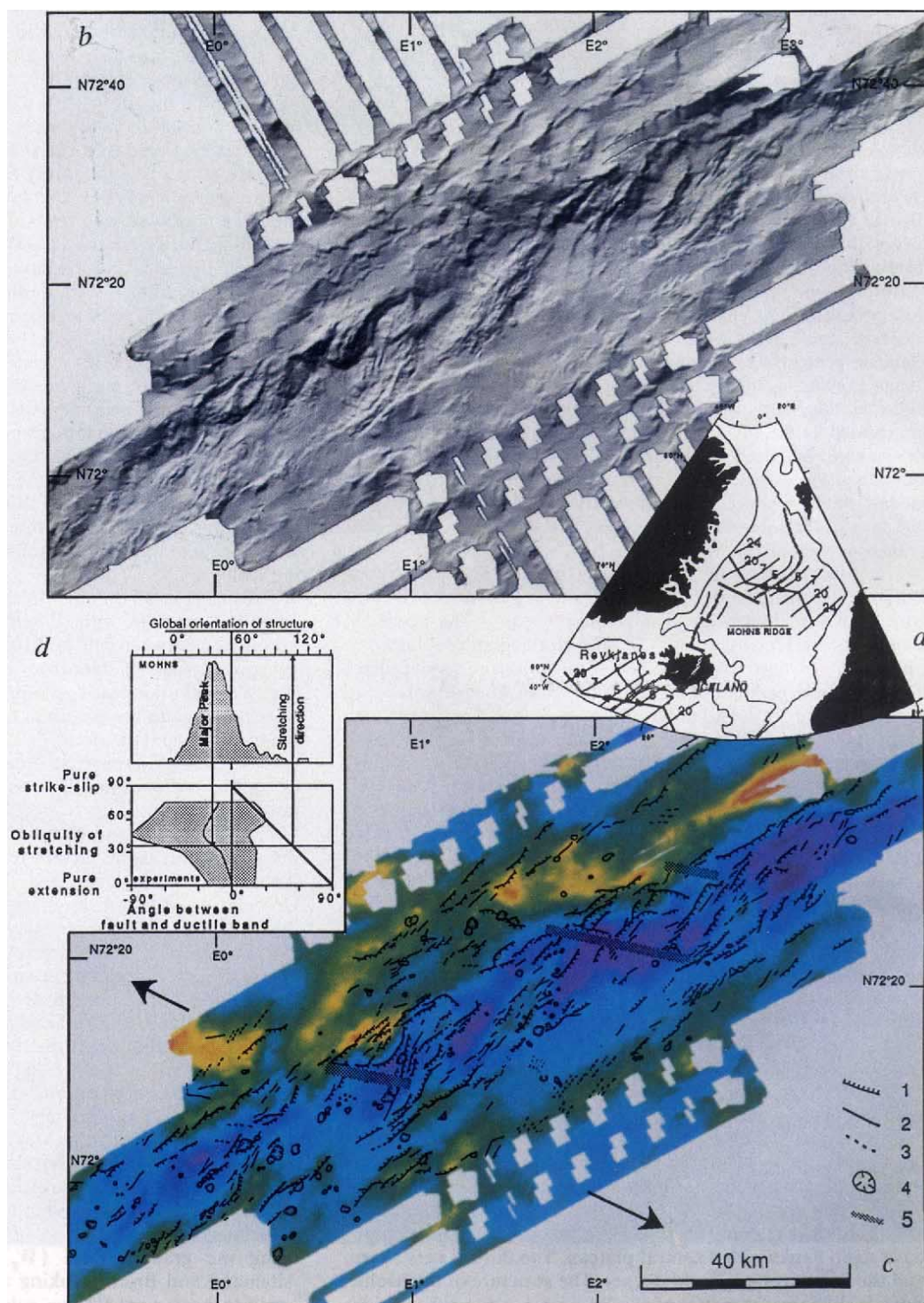
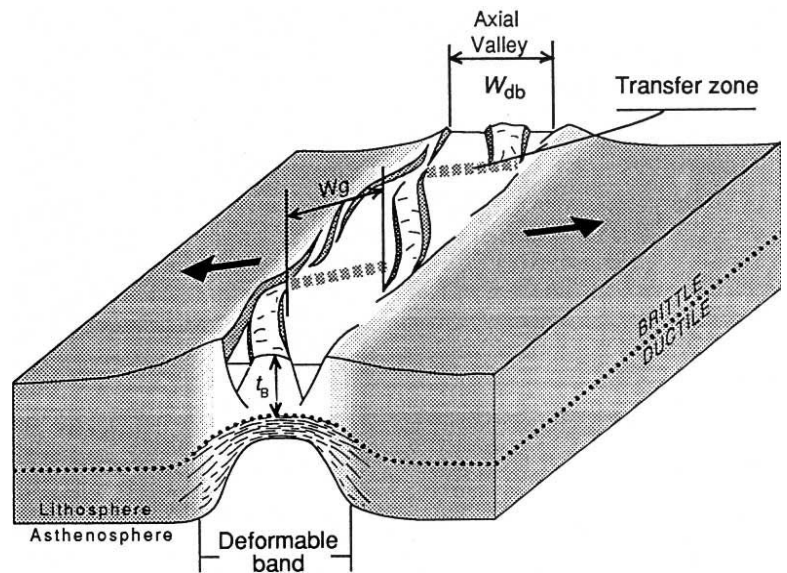


FIG. 2 Three-dimensional diagram of oblique rifting of the lithosphere at a mid-ocean ridge. Note the relations between the width of the deformable band (W_{db}) and the ductile layer. The thickness of the brittle layer (t_b) is estimated from the width of the graben (W_g) where $t_b = W_g/2 \tan 60^\circ$.



to the stretching direction: these bands correspond to the transfer zone identified in the Mohs ridge. With increasing extension, the ends of sigmoidal faults became parallel and interconnect, particularly along one border of the deformation zone where they produced a sharp boundary to the axial valley. This phenomenon increased with decreasing width-to-thickness ratio of the brittle layer. When the ratio of the width of the deformable band to the thickness of the brittle (W_{db}/t_b) was greater or equal to 3, the walls of the axial valley became straight.

Structural data obtained from Seabeam imagery confirm that the principal structural features of the Mohs ridge result from oblique rifting. A similar conclusion was reached by Searle and Laughton¹, who studied the Reykjanes ridge farther south, but we stress that the oblique rifting results from the sudden 50° rotation of the spreading direction 27 Myr ago. This interpretation implies that the present ridge corresponds to a deformable brittle-ductile band of oceanic lithosphere inherited from that period. The survival of oblique rifting for 27 Myr raises an

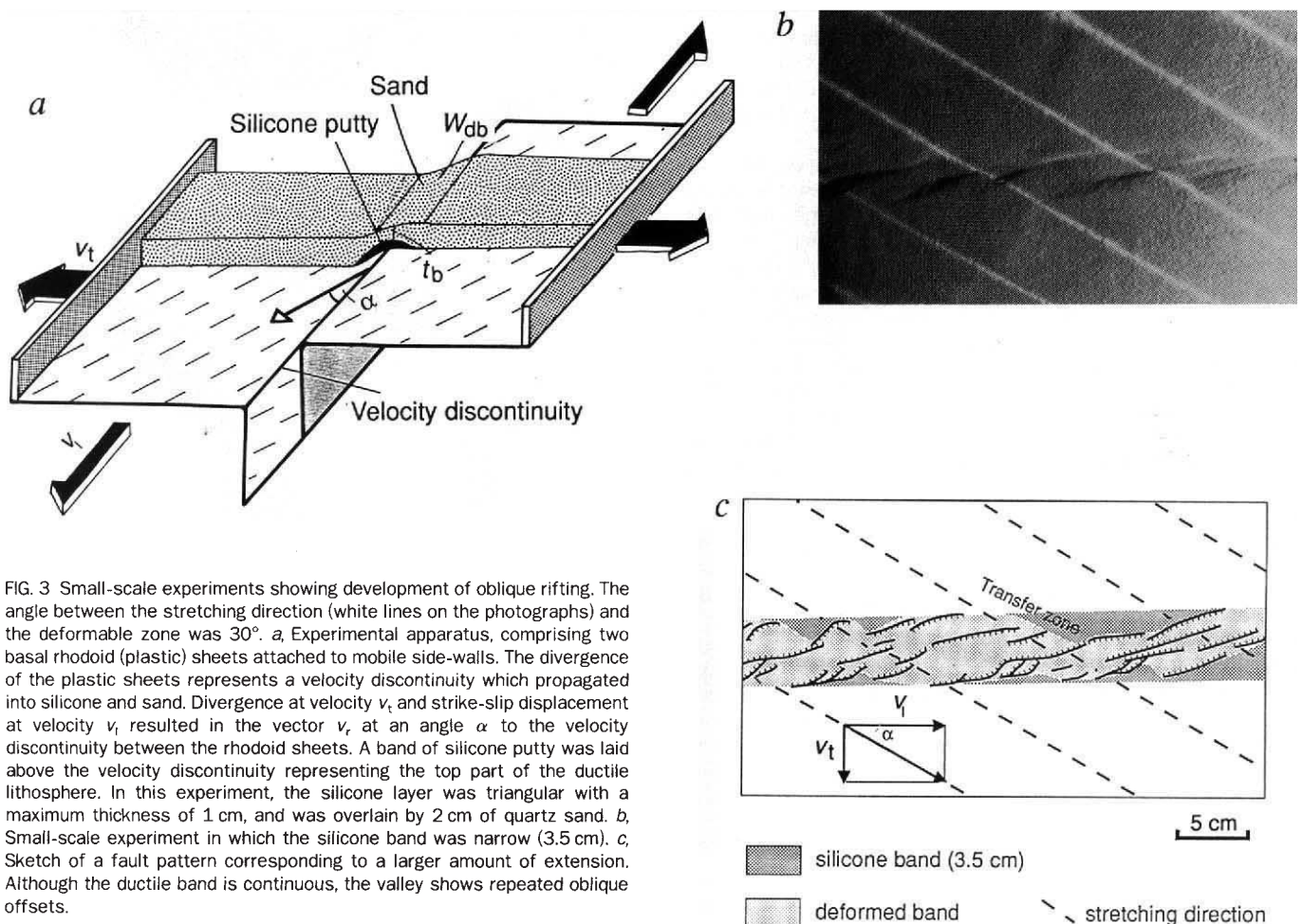


FIG. 3 Small-scale experiments showing development of oblique rifting. The angle between the stretching direction (white lines on the photographs) and the deformable zone was 30° . *a*, Experimental apparatus, comprising two basal rhodoid (plastic) sheets attached to mobile side-walls. The divergence of the plastic sheets represents a velocity discontinuity which propagated into silicone and sand. Divergence at velocity v_t and strike-slip displacement at velocity v_s resulted in the vector v_t at an angle α to the velocity discontinuity between the rhodoid sheets. A band of silicone putty was laid above the velocity discontinuity representing the top part of the ductile lithosphere. In this experiment, the silicone layer was triangular with a maximum thickness of 1 cm, and was overlain by 2 cm of quartz sand. *b*, Small-scale experiment in which the silicone band was narrow (3.5 cm). *c*, Sketch of a fault pattern corresponding to a larger amount of extension. Although the ductile band is continuous, the valley shows repeated oblique offsets.

unsolved thermomechanical problem. The answer may involve instabilities resulting from variations in spreading rates⁴. On the other hand, the system may be evolving towards a steady-state rift-transform configuration, in which the transfer zones will eventually develop into transform faults. □

Received 19 August; accepted 26 November 1992.

1. Searle, R. C. & Laughton, A. S. *Oceanol. Acta* **C4**, 5–13 (1981).
2. Renard, V. *et al. Terra Cogn.* **1**, 206 (1989).
3. Dauteuil, O., Brun, J.-P., Avedik, F. & Gely, L. *C.R. Acad. Sci. Paris* **311**, 357–363 (1990).
4. Geli, L. *C.R. Acad. Sci. Paris* **313**, 1291–1300 (1991).
5. Macdonald, K. C. *et al. Nature* **335**, 217–225 (1988).
6. Tron, V. & Brun, J.-P. *Tectonophysics* **188**, 71–84 (1991).
7. Vogt, P. R., Perry, R. K., Feden, R. H., Flemming, H. S. & Cherkis, N. A. *The Arctic Ocean*, Vol. 5, 493–598 (1981).
8. Vogt, P. R., Kovacs, L. C., Bernero, C. & Srivastava, S. P. *Tectonophysics* **89**, 95–160 (1982).
9. Talwani, M. & Eldholm, O. *Geol. Soc. Am. Bull.* **88**, 969–999 (1977).
10. Untermeier, P. thesis, Univ. de Bretagne Occidentale (1982).
11. Lecomte, I. thesis, Strasbourg Univ. (1991).
12. Gibbs, A. D. *J. geol. Soc. Lond.* **141**, 609–620 (1984).
13. Allemand, P. & Brun, J.-P. *Tectonophysics* **188**, 63–69 (1991).

ACKNOWLEDGEMENTS. We thank F. Avedik, L. Geli and V. Renard for the Seabeam data and for discussions, and N. Arndt for advice.

Magnitude and patterns of herbivory in aquatic and terrestrial ecosystems

Hélène Cyr^{*†‡} & Michael L. Pace[†]

[†] Institute of Ecosystem Studies, Box AB, Millbrook, New York, 12545 USA

[‡] Ecology Program, Rutgers University, Piscataway, New Jersey 08855-1059, USA

HERBIVORES can consume a sufficiently large proportion of primary production to regulate plant biomass in some environments^{1–3}. Little is known, however, about how rates of herbivory vary among ecosystems and how herbivores influence the global distribution of vegetation. Patterns of herbivory in terrestrial ecosystems have been summarized recently^{4,5}, but comparisons with aquatic systems are uncertain because past generalizations about herbivory in aquatic systems are based on surprisingly few data^{6–8}. Herbivory is thought to be higher in aquatic than in terrestrial ecosystems^{9–11} and the impact of herbivores in aquatic systems is believed to decrease with increasing primary productivity^{12–15}, a pattern opposite to that in terrestrial systems^{4,5}. Here we examine these hypotheses using data from 44 aquatic sites. Herbivore biomass and herbivory rates increase at similar rates with increasing primary productivity in aquatic and in terrestrial systems. For a given level of primary productivity, aquatic and terrestrial herbivores reach similar biomass, but aquatic herbivores remove on average 51% of annual primary production, three times more than terrestrial herbivores. Mass-specific rates of herbivory are greater in aquatic than in terrestrial systems.

To compare the impact of herbivores in different types of ecosystems, we compiled published measurements of herbivory rates, herbivore biomass and primary productivity for a wide range of aquatic systems^{16–38} and used the data of McNaughton *et al.*^{4,5} for terrestrial systems. A modified version of this terrestrial data set³⁹, which condenses data from different years into long-term averages and excludes sites where the herbivores were managed, is slightly different, but does not alter the comparisons we make between aquatic and terrestrial systems. Herbivory is defined as ingestion of plant material, not its assimilation, and includes the effect of all major herbivores. Studies were discarded when herbivory and/or primary productivity were estimated indirectly, measured over a short time that could not be extrapo-

lated to annual values, or were not measured concurrently (for example, rates of herbivory and primary productivity measured in different years). Units were standardized, preferably with conversion factors measured by the original authors but if these were not available, we used standard values^{31,40–43}. Copies of the data set are available from the authors.

Herbivory rates vary among types of ecosystems (Fig. 1). Algae are grazed more intensively than aquatic vascular plants (Fig. 1a, b) and herbivores generally remove a larger proportion of net primary production in aquatic (median, 51%; $n = 44$) than in terrestrial systems (median, 18%). These results support the hypothesis that herbivory is generally greater in aquatic than in terrestrial ecosystems^{9–11}, even for aquatic macrophytes, which were traditionally thought to be largely immune to herbivory⁴⁴.

Herbivory rates increase with increasing annual net primary productivity. This relationship in aquatic systems is linear (slope of log-log relationship is not different from 1, t -test: $P = 0.7$; Fig. 2), suggesting that herbivores remove the same proportion of annual net primary production in poor as in rich aquatic systems. Similar relationships are found for algae and for aquatic vascular plants (comparison of slopes using an F -test: $P = 0.7$, after removal of the St Lawrence site (Fig. 2) to equalize the variance between the two groups of data), although algae are grazed 2.5 times more than vascular plants (ANCOVA,

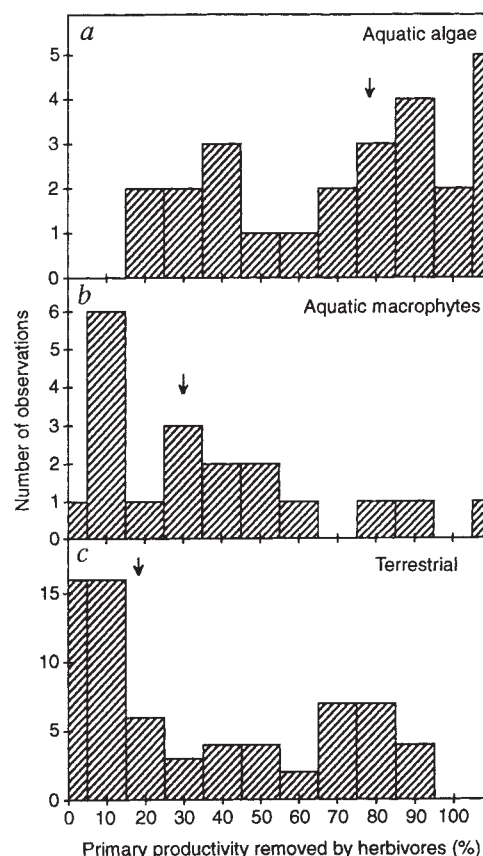


FIG. 1 Frequency distributions of the proportion of annual net primary productivity removed by herbivores in a, aquatic algae (phytoplankton^{16–28}, $n = 17$, and reef periphyton^{29,30}, $n = 8$); b, submerged^{31–35} ($n = 5$) and emergent^{36–38} ($n = 14$) vascular plants; and c, terrestrial plants^{4,5} ($n = 67$). Primary productivity and herbivory data for aquatic and terrestrial rooted plants are limited to the aboveground portion of the plants. Comparison of herbivory on algae and rooted plants, therefore, assumes that the proportion of aboveground primary productivity grazed is representative of the whole plant, an assumption supported by the limited data available from terrestrial systems (aboveground^{4,5}: median, 18%, $n = 69$; belowground^{50–52}: median, 13%, $n = 14$). Arrows indicate median values (aquatic algae, 79%; aquatic macrophytes, 30%; terrestrial plants, 18%).

* Present address: Department of Biology, McGill University, 1205 Docteur Penfield, Montréal, Québec H3A 1B1, Canada.

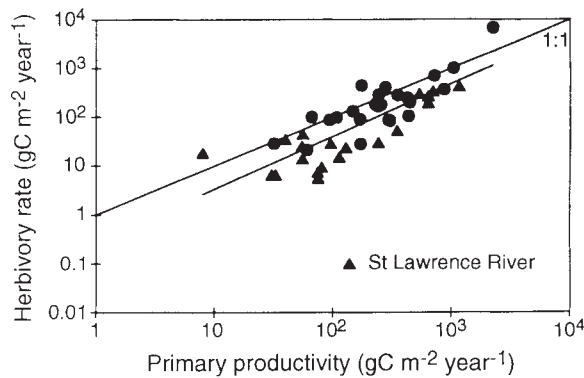


FIG. 2 Relationship between annual rates of herbivory (HR) and net primary productivity (PP) in aquatic systems ($\log_{10} HR = 1.07 \log_{10} PP - 0.55$; $r^2 = 0.45$, $n = 44$; removing low St Lawrence site: $\log_{10} HR = 1.03 \log_{10} PP - 0.39$; $r^2 = 0.63$). Circles represent planktonic and periphytic algae, triangles represent submerged and emergent vascular plants. The 1:1 line shows where 100% of primary production is grazed.

$P < 0.01$). These results do not support reports that herbivory declines along gradients of primary productivity in marine and freshwater plankton^{12-15,45}. Our analysis is more general than previous comparisons of aquatic systems^{12,15,45} because it is based on more data ($n = 44$, compared with $n = 3-9$), and covers the whole year rather than short sampling periods (one day to three months). The pattern we found in aquatic systems is not significantly different from the pattern in terrestrial systems (comparison of slopes using an F -test: $P > 0.6$; Fig. 3).

Higher rates of herbivory in aquatic systems (Fig. 3) suggest two non-exclusive corollaries: herbivores maintain a higher biomass and/or graze more per unit of biomass in aquatic compared with terrestrial systems. We compared annual average herbivore biomass among aquatic systems. Herbivore biomass was reported or could be calculated for 30 sites (Fig. 4). In two submerged macrophyte beds, herbivores reached much higher average biomass than in other aquatic systems of similar productivity (upper two triangles in Fig. 4; refs 32, 33). Excluding these two sites, average herbivore biomass in aquatic systems increases with net primary productivity (ANOVA: $P < 0.001$, $r^2 = 0.57$, $n = 28$; Fig. 4, filled symbols). The strength of this relationship may be surprising as some sites support mostly resident herbivores (planktonic invertebrates, coral reef fish and invertebrates; filled circles in Fig. 4), whereas others are mostly grazed by migratory herbivores (waterfowl, turtles; triangles in Fig. 4). The biomass of migratory herbivores might be expected to be poorly related to the productivity of a single feeding site, but the pattern in biomass of migratory herbivores agrees well with that of resident herbivores.

Annual average herbivore biomass also increases with increasing net primary productivity in terrestrial systems^{4,5} (Fig. 4, open symbols). McNaughton *et al.* report a slope of 1.52 ± 0.18 (s.e.),

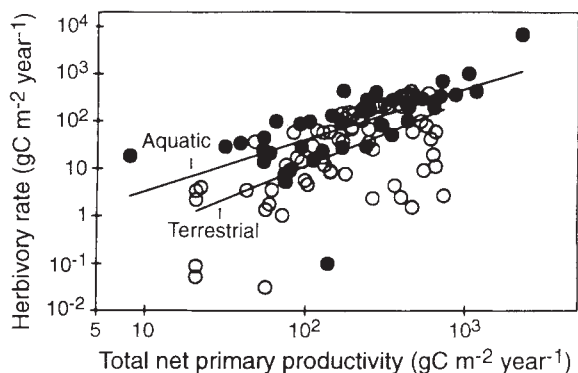


FIG. 3 Relationship between annual rates of herbivory and net primary productivity in aquatic (filled circles; slope $b \pm \text{s.e.} = 1.07 \pm 0.18$) and terrestrial (open circles; $b \pm \text{s.e.} = 1.38 \pm 0.22$) ecosystems. Rates of herbivory are on average three times higher in aquatic than in terrestrial systems (ANCOVA: $P < 0.001$).

but their relationship is strongly influenced by a tundra site with very low herbivore biomass (open circle with lowest herbivore biomass in Fig. 4). When this point is removed from the analysis, which also equalizes the variance of the aquatic and the terrestrial data, the relationship in terrestrial systems resembles that in aquatic systems (comparison of slopes using an F -test: $P = 0.3$; ANCOVA, $P = 0.15$; common slope: $b \pm \text{s.e.} = 1.17 \pm 0.11$, $r^2 = 0.60$, $n = 78$). For a given level of net primary productivity, herbivores reach similar average biomass in aquatic and terrestrial ecosystems.

Higher herbivory rates in aquatic systems (Fig. 3), despite similar herbivore biomass (Fig. 4), suggest that mass-specific herbivory rates are higher in aquatic than in terrestrial systems. Differences in mass-specific herbivory rates have been attributed to characteristics of the plants (such as nutrient content, chemical or physical defences⁴⁶⁻⁴⁸) and/or the herbivores (such as size, ectothermy versus endothermy, behaviour⁴⁶⁻⁴⁸), but the importance of these and other factors in determining herbivory at the scale of ecosystems remains to be tested.

Similarities and differences in herbivory have important implications for patterns of energy flow in ecosystems. Higher herbivory rates in aquatic than in terrestrial systems suggest that herbivores in aquatic systems recycle limiting nutrients more rapidly to primary producers and leave a smaller proportion of the plant material produced annually to decomposers. Higher herbivory rates in aquatic systems also suggest higher herbivore productivity because production efficiencies (ratio herbivore productivity: herbivory rate) do not differ systematically

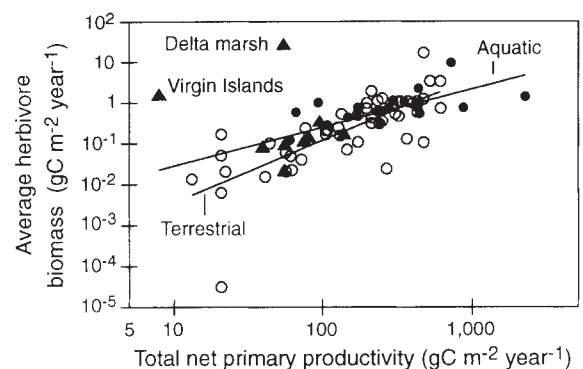


FIG. 4 Relationship between annual average herbivore biomass (HB) and net primary productivity (PP) in aquatic and terrestrial systems. Filled symbols represent aquatic sites (circles for algae¹⁶⁻³⁰; triangles for vascular plants³¹⁻³⁸); open symbols represent terrestrial sites^{4,5}. Waterfowl densities were transformed to biomass using species-specific average individual mass⁵³, assuming a 1:1 sex ratio. The relationship shown for aquatic systems excludes two submerged macrophyte beds (Delta marsh and Virgin Island excluded; $\log_{10} HB_{\text{aquatic}} = 0.94 \log_{10} PP - 2.48$; $r^2 = 0.57$, $n = 28$) and the relationship shown for terrestrial systems excludes one tundra site ($\log_{10} HB_{\text{terrestrial}} = 1.28 \log_{10} PP - 3.38$; $r^2 = 0.61$, $n = 50$).

between aquatic and terrestrial organisms⁴⁹. Assuming that herbivore productivity is higher in aquatic than in terrestrial systems, and that there is no systematic difference in herbivore biomass (Fig. 4), immigration, emigration and non-predatory mortality among types of ecosystems, we expect predation to be higher on aquatic than on terrestrial herbivores. □

Received 10 August; accepted 27 October 1992.

- Power, M. E., Stewart, A. J. & Matthews, W. J. *Ecology* **69**, 1894–1898 (1988).
- Feminella, J. W. & Resh, V. H. *Holarctic Ecol.* **12**, 1–8 (1989).
- Keats, D. W., South, G. R. & D. H. Steele, *Mar. Ecol. Prog. Ser.* **68**, 181–193 (1990).
- McNaughton, S. J., Oesterheld, M., Frank, D. A. & Williams, K. J. *Nature* **341**, 142–144 (1989).
- McNaughton, S. J., Oesterheld, M., Frank, D. A. & Williams, K. J. in *Comparative Analyses of Ecosystems* (eds Cole, J., Lovett, G. & Findlay, S.) 120–139 (Springer, New York, 1991).
- Wiegert, R. G. & Owen, D. F. *J. theor. Biol.* **30**, 69–81 (1971).
- Whittaker, R. H. & Likens, G. E. in *Carbon and the Biosphere* (eds Woodwell, G. M. & Pecan, E. V.) 281–300 (Proc. 24th Brookhaven Symp. Biol., US Atomic Energy Commission, 1973).
- Pimentel, D., Levin, S. A. & Soans, A. B. *Ecology* **56**, 381–390 (1975).
- Whittaker, R. H. *Communities and Ecosystems* 2nd edn (Macmillan, New York, 1975).
- Valliel, I. *Marine Ecological Processes* (Springer, New York, 1984).
- Ricklefs, R. *Ecology*, 3rd edn (Freeman, New York, 1990).
- Glizwicz, Z. M. & Hillbricht-likowska, A. *Verh. int. Verein. theor. angew. Limnol.* **18**, 197–203 (1972).
- Eppley, R. W. in *Analysis of Marine Ecosystems* (ed. Longhurst, A. R.) 343–361 (Academic, London, 1981).
- McQueen, D. J., Post, J. R. & Mills, E. L. *Can. J. Fish. aquat. Sci.* **43**, 1571–1581 (1986).
- Sager, P. E. & Richman, S. *Can. J. Fish. aquat. Sci.* **48**, 116–122 (1991).
- Gophen, M. *J. Plank. Res.* **3**, 15–24 (1981).
- Gulati, R. D., Siewertsen, K. & Postema, G. *Hydrobiologia* **95**, 127–163 (1982).
- Déssery, S., Lancelot, C. & Billen, G. *Verh. int. Verein. theor. angew. Limnol.* **22**, 1504–1509 (1984).
- Gulati, R. D. *Verh. int. Verein. theor. angew. Limnol.* **22**, 863–867 (1984).
- Gulati, R. D. *Hydrobiologia* **200/201**, 99–118 (1990).
- Welschmeyer, N. A. & Lorenzen, C. J. *Limnol. Oceanogr.* **30**, 1–21 (1985).
- Hart, R. C. *J. Plank. Res.* **8**, 401–426 (1986).
- Jarvis, A. C. *J. Plank. Res.* **8**, 1065–1078 (1986).
- Jespersen, A. M., Christoffersen, K. & Riemann, B. *Verh. int. Verein. theor. angew. Limnol.* **23**, 440–444 (1988).
- Van Donk, E., Grimm, M. P., Gulati, R. D. & Klein Breteler, J. P. G. *Hydrobiologia* **200/201**, 275–289 (1990).
- Van Donk, E. *et al. Hydrobiologia* **200/201**, 291–301 (1990).
- Garnier, J. & Mourelatos, S. *Freshwater Biol.* **25**, 515–523 (1991).
- Pinto-Coelho, R. M. *Verh. int. Verein. theor. angew. Limnol.* **24**, 842–845 (1991).
- Hatcher, B. G. *Proc. 4th Int. Coral Reef Symp.* **2**, 515–524 (1981).
- Carpenter, R. C. *Ecol. Monogr.* **56**, 345–363 (1986).
- Jupp, B. P. & Spence, D. H. N. *J. Ecol.* **65**, 431–446 (1977).
- Anderson, M. G. & Low, J. B. *J. Wildl. Mgmt.* **40**, 233–242 (1976).
- Williams, S. L. *Marine Biol.* **98**, 447–455 (1988).
- Greenway, M. *Aquat. Bot.* **2**, 117–126 (1976).
- Verhoeven, J. T. A. *Aquat. Bot.* **8**, 209–253 (1980).
- Cargill, S. M. & Jeffries, R. L. *J. appl. Ecol.* **21**, 669–686 (1984).
- Smith, L. M. & Kadlec, J. A. *Ecology* **66**, 259–265 (1985).
- Giroux, J. F. & Bédard, J. *J. appl. Ecol.* **24**, 773–788 (1987).
- Moen, J. & Oksanen, L. *Nature* **353**, 510 (1991).
- Boyd, C. E. in *Freshwater Wetlands* (eds Good, R. E., Whigham, D. F. & Simpson, R. L.) 155–167 (Academic, New York, 1978).
- DeNoyelles, F. & Likens, G. E. in *An Ecosystem Approach to Aquatic Ecology* (ed. Likens, G. E.) 161–175 (Springer, New York, 1985).
- Behrendt, H. *Arch. Hydrobiol.* **118**, 129–145 (1990).
- Duarte, C. M. *Mar. Ecol. Prog. Ser.* **67**, 201–207 (1990).
- Lodge, D. M. *Aquat. Bot.* **41**, 195–224 (1991).
- Elser, J. J. & Goldman, C. R. *Limnol. oceanogr.* **36**, 64–90 (1991).
- Lubchenco, J. & Gaines, S. D. *A. Rev. ecol. Syst.* **12**, 405–437 (1981).
- Crawley, M. J. *Herbivory* (Blackwell, Oxford, 1983).
- Hay, M. E. & Fenical, W. A. *Rev. ecol. Syst.* **19**, 111–145 (1988).
- Schroeder, L. A. *J. theor. Biol.* **93**, 805–828 (1981).
- Scott, J. A., French, N. R. & Leatham, J. W. in *Perspectives in Grassland Ecology* (ed. French, N. R.) 89–105 (Springer, New York, 1979).
- Anderson, D. C. & MacMahon, J. A. *Ecol. Monogr.* **51**, 179–202 (1981).
- Ingham, R. E. & Detling, J. K. *Oecologia* **63**, 307–313 (1984).
- Kortright, F. H. *The Ducks, Geese and Swans of North America* (Stackpole Company, Harrisburg, PA, and Wildlife Management Institute, Washington DC, 1967).

ACKNOWLEDGEMENTS. We thank S. McNaughton and his group, and J. Moen and L. Oksanen for sharing data, and G. Likens, R. Ostfeld and R. Peters for their comments. HC was supported by an FCAR scholarship from the Québec Government and the graduate school of Rutgers University. Funding was from NSF.

Preferential use of organic nitrogen for growth by a non-mycorrhizal arctic sedge

F. Stuart Chapin III*, Lori Moilanen† & Knut Kielland†

* Department of Integrative Biology, University of California, Berkeley, California 94720, USA

† Institute of Arctic Biology, University of Alaska, Fairbanks, Alaska 99775, USA

PLANT growth in arctic tundra is strongly nitrogen-limited despite large pools of soil organic nitrogen^{1–4}. Here we report that field-collected roots of *Eriophorum vaginatum*, an arctic sedge, rapidly absorb free amino acids, accounting for at least 60% of the nitrogen absorbed by this species in the field. In solution culture, *Eriophorum* accumulates more nitrogen and biomass when supplied with amino acids than when grown on inorganic nitrogen, whereas *Hordeum vulgare* (a cereal adapted to mineral soils) grows least when nitrogen is supplied as amino acids. To our knowledge, this is the first documentation of preferential absorption and use of organic nitrogen by a non-mycorrhizal vascular plant. The direct absorption of amino acids by *Eriophorum* short-circuits the bottleneck in arctic nitrogen cycles imposed by temperature-limited mineralization.

In the moist tundra of arctic Alaska, mineralization of nitrogen from organic to inorganic form accounts for only about half of the annual nitrogen requirement^{1,2} of this nitrogen-limited^{2–4} vegetation. How do plants meet the rest of their nitrogen requirement? In well-drained heaths, ericoid mycorrhizal fungi (and perhaps ectomycorrhizal fungi of deciduous shrubs) hydrolyse proteins and transfer amino acids to their symbiotic host plants^{5–9}. But moist and wet tundras, which account for 40% of the area of the Low Arctic¹⁰, are dominated by sedges that are generally non-mycorrhizal or have an endomycorrhizal symbiont^{5–7,11} that cannot hydrolyse organic N¹¹. Moist and wet

tundra soils have higher concentrations of water-extractable free amino acids (2–8 µg N g⁻¹ dry soil) than of inorganic nitrogen (0.5–1.1 µg N g⁻¹) (ref. 2), perhaps because protease activity is an order of magnitude faster than net nitrogen mineralization¹². Here we demonstrate that *Eriophorum vaginatum* L., a non-mycorrhizal sedge¹² that dominates moist upland tundra throughout the circumpolar Arctic, can absorb amino acids and grow on them as its sole nitrogen source. By contrast, barley (*Hordeum vulgare* L., cv. Stepto), a north-temperate cereal that grows on mineral soils, does not efficiently utilize amino acids for growth.

In upland moist tundra at Toolik Lake in arctic Alaska (68° N, 149° W; 760 m elevation) we collected *Eriophorum* roots and measured kinetics of uptake of ¹⁴C-labelled methylamine (an ammonium analogue), glycine, glutamate and aspartate in four separate experiments. *Eriophorum* had a higher capacity ($V_{\max}$) for ammonium than for amino-acid uptake, but higher affinity (lower apparent K_M) for most amino acids (Table 1). On the basis of these kinetics and the seasonal average concentration of free amino acids in the soil solution, we estimate that the field absorption rate of ammonium was lower than that of glycine but higher than that of aspartate or glutamate (Table 1). Assuming conservatively that other free amino acids that were present in the soil solution but omitted from our study² are not absorbed, amino acids would comprise 60% of the nitrogen absorbed by *Eriophorum*. Plants do absorb other amino acids^{13,14}, so amino acids probably comprise an even higher proportion of the nitrogen acquired by *Eriophorum* in the field. The rates of methylamine uptake in our experiments were similar to ammonium uptake rates reported for *Eriophorum* using ¹⁵N-ammonium, or to net uptake by intact plants measured with specific ion electrodes^{15–18}; this indicates that our uptake rates are not an artefact resulting from the use of methylamine as an ammonium analogue.

When grown in solution culture, *Eriophorum* accumulated most nitrogen and biomass when supplied with amino acids as its sole nitrogen source (Fig. 1). By contrast, *Hordeum*, a species that typically grows on mineral soils, grew most on nitrate and

ammonium, achieving nearly twice the biomass as when grown on amino acids. The form of nitrogen supplied to plants had no effect on root-to-shoot ratio in either species, indicating that species differences in capacity to use organic rather than inorganic nitrogen involves physiological properties of roots rather than differences in biomass allocation. When grown on nitrate, *Eriophorum* had a lower capacity to reduce nitrate *in vitro* ($0.8 \pm 0.1 \mu\text{mol g}^{-1} \text{h}^{-1}$ in the shoot, and $14.6 \pm 1.4 \mu\text{mol g}^{-1} \text{h}^{-1}$ in the root) than did *Hordeum* ($131 \pm 9 \mu\text{mol g}^{-1} \text{h}^{-1}$ in the shoot and $275 \pm 18 \mu\text{mol g}^{-1} \text{h}^{-1}$ in the root on a dry-mass basis¹⁹), suggesting that low levels of nitrate reductase may partially explain the slow growth of *Eriophorum* on nitrate. This may reflect low nitrate availability in the moist tundra soils where *Eriophorum* evolved^{2,12,20}. Neither species had measurable nitrate reductase activity when grown on nitrogen sources other than nitrate.

Growth in amino acids had no effect on bacterial numbers in *Eriophorum* hydroponic solutions but caused a 35-fold increase in bacteria in solutions in which *Hordeum* was growing (Fig. 2). These results are consistent with the growth experiments, indicating that *Eriophorum* is more effective than *Hordeum* in removing amino acids from nutrient solutions in the face of bacterial competition.

Because we measured the carbon (^{14}C) moiety in field uptake experiments and because *Eriophorum* grew better on amino acids than on ammonium, plants probably absorbed amino acids intact rather than as ammonium after deamination by bacteria in the growth solution. This conclusion was supported by changes in solution pH during the growth experiment. Two of the amino acids in the growth solution (glutamate and aspartate) are strong anions at the solution pH of 5.5, two are dipolar neutral ions (alanine and glycine; isoelectric point pH 5–6), and one is a cation (arginine). During the last week of the experiment, when uptake was most rapid, plants supplied with ammonium reduced the pH (Fig. 2) as a result of proton secretion to balance

TABLE 1 Seasonal average concentration of ammonium and free amino acids in tussock tundra and kinetics of uptake

N form	Field concentration (μM)	V_{max} ($\mu\text{mol g}^{-1} \text{h}^{-1}$)	K_{M} (μM)	Predicted field uptake rate ($\mu\text{mol g}^{-1} \text{h}^{-1}$)
Ammonium	10.7	13.7	242	0.58
Glycine	8.4	2.0	12	0.82
Aspartate	7.9	4.1	293	0.11
Glutamate	5.2	0.6	118	0.03
Other amino acids	84.1	—	—	—

Ammonium and amino-acid concentrations were measured by HPLC²⁵ on a water extract immediately after collecting the upper 10 cm of organic soil in June, July and August ($n=8$ per sample date). Roots of *Eriophorum vaginatum* were gently removed from the peat and brought to the field laboratory in 0.5 mM CaCl_2 . Because *Eriophorum* roots are unbranched and the peat in which they grow is loose (bulk density, 0.17 g cm^{-3})¹², they can be collected without detectable damage and yield physiologically meaningful uptake kinetics^{15–17}. Roots were placed in one of five concentrations (10–500 μM) of ^{14}C -labelled methylamine, glycine, glutamate or aspartate at 14°C (maximum July soil temperature at 5–10 cm depth) for 10 min, rinsed for 2 min in 1 mM KCl to exchange any unabsorbed amino acids, dried, weighed, oxidized, and the trapped CO_2 counted by liquid scintillation²⁶. Rates are expressed on a root dry-mass basis. All uptake and rinse solutions contained 0.5 mM CaCl_2 to maintain membrane integrity. Kinetic parameters for each compound were calculated from four measurements per concentration, assuming Michaelis–Menton kinetics²⁷.

excess cation uptake²¹. By contrast, plants grown in nitrate and amino acids caused pH to rise, indicating that these forms are absorbed as anions balanced by proton absorption. If amino acids had been hydrolysed before uptake and nitrogen absorbed as ammonium, then ammonium uptake would have caused a decrease in pH.

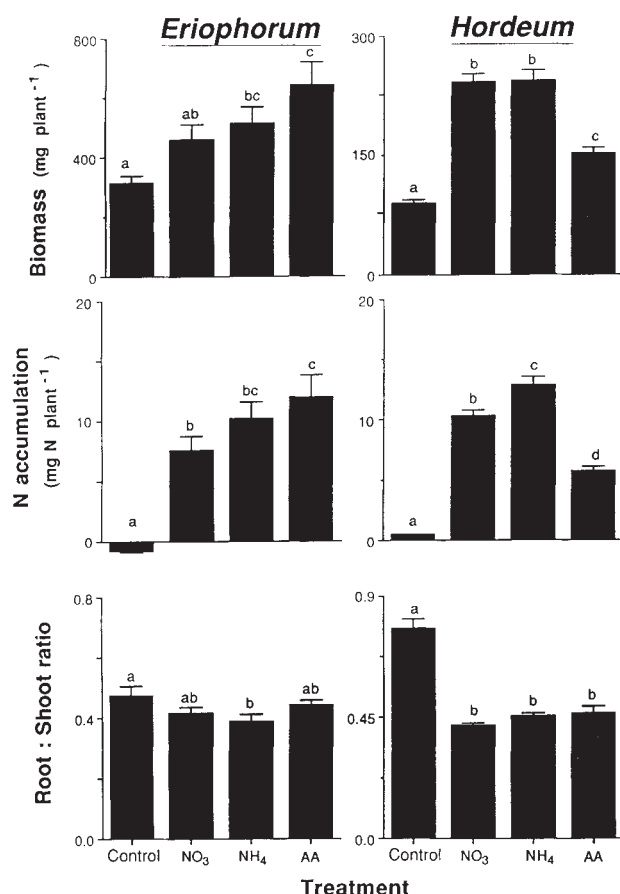


FIG. 1 Final biomass, nitrogen accumulation during the experiment, and ratio of root biomass to shoot biomass at harvest of *Eriophorum* plants grown for 24 d, and *Hordeum* plants grown for 11 d, in one of four treatments: zero N (control), nitrate, ammonium, or amino acids (AA). Data are means $\pm$ standard error, $n=15$. Treatments associated with the same letter (a, b, c above bars) are not significantly different ($P < 0.05$).

METHODS. Seeds of *Hordeum vulgare* L. cv. Stepto were surface-sterilized with 0.5% NaOCl, rinsed, germinated on wet cheesecloth in distilled water in a growth chamber with a 20°C 20 h day and a 15°C 4 h night. After one week, seedlings were transferred to 1.8-litre containers of aerated half-strength Hoagland's solution containing 1.0 mM nitrogen and 10 p.p.m. ampicillin, an antibiotic that at this concentration causes no visible damage to roots and minimizes bacterial growth. We found that plants grown in sterile culture had small-diameter finely branched roots, unlike the thick unbranched roots observed in the field and in the non-sterile solution culture used in our experiments. The amino-acid treatment had equimolar concentrations of glycine, glutamic acid, aspartic acid, alanine and arginine, amino acids that are abundant in Alaskan tundra soils². Solutions were adjusted to pH 5.5, which is near-optimal for growth of both barley (pH 6.0) and tundra sedges (pH 5.0). Solutions were changed daily, at which time we measured solution pH. *Eriophorum* seeds from Toolik Lake were sown and grown for a month on a peat-vermiculite mix. Seedlings were then rinsed free of soil and transplanted into hydroponic solutions in the growth chamber. Conditions were identical to those of the barley experiment except that we used less concentrated Hoagland's solution (0.1 strength) and nitrogen concentration (0.1 mM). Preliminary experiments showed that, when *Eriophorum* seedlings were grown in 0.25-strength Hoagland's solution, they excreted salts on the leaves, had substantial root death, and grew slowly. Therefore, we could find no single growth solution that provided optimal growth for both *Hordeum* and *Eriophorum*. Plants were harvested after 24 d, separated into roots and shoots, dried, weighed, ground in a Wiley mill, digested in selenous sulphuric acid and analysed for nitrogen (nitroferrocyanide procedure)²⁸.

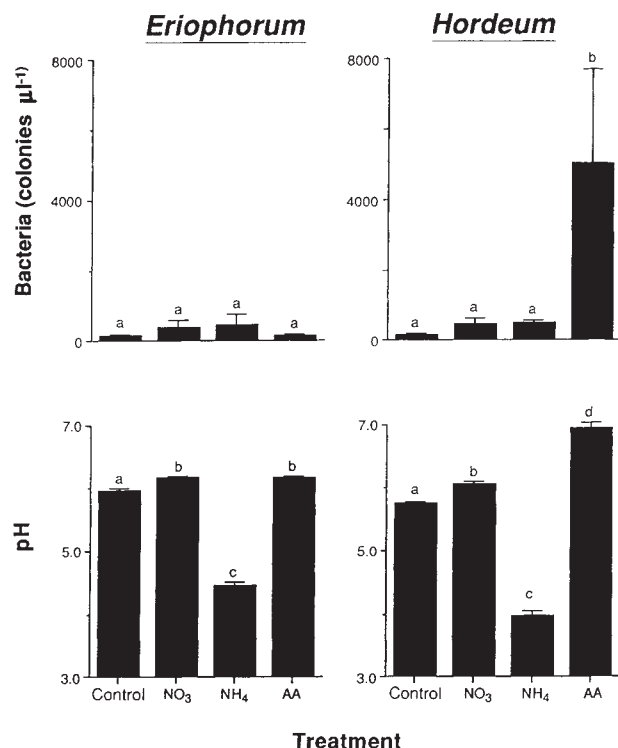


FIG. 2 Bacterial density and pH of nutrient solutions 24 h after replenishment, averaged over the last week of the experiment. Plants of *Eriophorum* and *Hordeum* were grown in each of four nitrogen treatments: zero N (control), nitrate, ammonium, or amino acids (AA). Statistics as in Fig. 1. Viable bacterial counts in the hydroponic solutions were determined by dilution plating on Difco tryptic soy agar²⁹. Dilutions of 10^{-4} -fold gave final colony counts of 30–300 colonies per plate. Inoculated plates were inverted and incubated for 2–4 d at room temperature. Colonies were counted with a Quebec darkfield colony counter. Storage containers and hydroponic containers were washed daily with scalding hot water and rinsed with distilled water to minimize bacterial growth.

To confirm that plants of *Eriophorum* absorbed amino acids intact, we measured uptake rates of ^{14}C -glycine and of nitrate and ammonium at the end of the growth experiment. Plants grown on amino acids absorbed ^{14}C -labelled glycine more rapidly than those grown on inorganic nitrogen, whereas plants grown on inorganic nitrogen absorbed ammonium more rapidly than those grown on amino acids (Fig. 3). Plants in all treatments showed a high capacity to absorb nitrate (Fig. 3), but this did not translate into rapid growth on nitrate (Fig. 1), perhaps because of the low nitrate assimilation²². The major conclusion from the short-term uptake experiment is that plants growing in amino acids absorb glycine intact at substantial rates.

Plants and plant cells absorb amino acids at low rates in the absence of mycorrhizae^{13,14}, but to our knowledge this is the first study to demonstrate that a species adapts to organic soils through an enhanced capacity to absorb and utilize amino acids as a nitrogen source for growth. By directly absorbing and utilizing amino acids, *Eriophorum* acquires an energetically favourable source of nitrogen and short-circuits the mineralization process that is considered to be the bottleneck in the nitrogen cycle of northern ecosystems²⁰. This direct absorption of organic nitrogen may be widespread among arctic tundra species² and among plants with ericoid and ectomycorrhizal symbionts^{5–9}. *Eriophorum* roots also have an external phosphatase that mineralizes organic phosphate in soils and could supply 25–70 per cent of the plant's requirement²³. Thus, in tundra, plants influence nutrient cycles not only through effects on litter quality²⁴ and nutrient depletion by uptake¹², but also

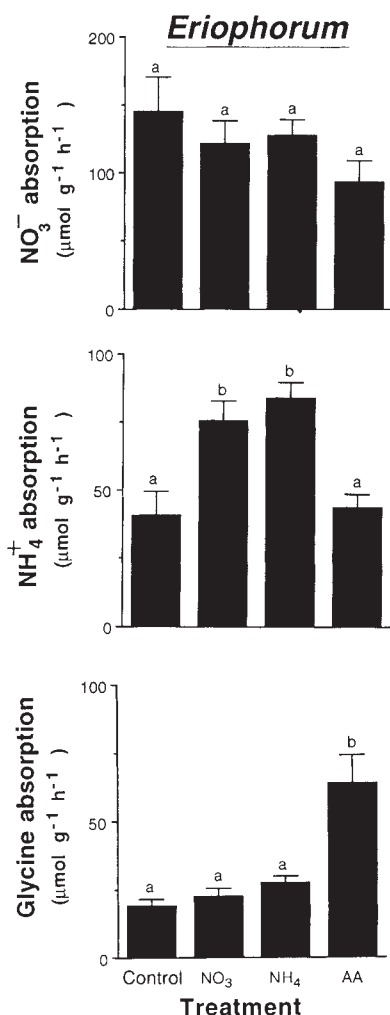


FIG. 3 Short-term measurements of absorption of nitrate, ammonium, and glycine by plants of *Eriophorum* grown in each of four nitrogen treatments: zero N (control), nitrate, ammonium, or amino acids (AA). Statistics as in Fig. 1. At the time of harvest, individual plants were rinsed in a solution of 0.5 mM CaCl_2 for 30 min, then transferred to 100 ml of 0.1 mM NH_4NO_3 , 0.1 mM ^{14}C -glycine, 0.5 mM CaCl_2 , so that uptake of nitrate, ammonium and an amino acid could be compared. Glycine was selected as the amino acid because it is absorbed rapidly by plants and is abundant in tundra soils (Table 1). Uptake was measured by solution depletion after 1.5 h. Control flasks without plants showed no change in nitrate, ammonium or glycine during the uptake period. Nitrate and ammonium were measured with specific ion electrodes³⁰ and ^{14}C -glycine by liquid scintillation.

by actively utilizing organic sources of nitrogen and phosphorus to accelerate nutrient turnover. In this ecosystem plants could directly control nutrient cycling to a greater extent than previously appreciated. □

Received 25 September; accepted 2 November 1992.

- Shaver, G. R., Nadelhoffer, K. J. & Giblin, A. E. in *Quantitative Methods in Landscape Ecology* (eds Turner, M. G. & Gardner, R. H.) 105–125 (Springer, New York, 1990).
- Kielland, K. *Processes Controlling Nitrogen Release and Turnover in Arctic Tundra* thesis, Univ. Alaska (1990).
- Haag, R. W. *Can. J. Bot.* **52**, 103–116 (1974).
- Shaver, G. R., Chapin III, F. S. & Gartner, B. L. *J. Ecol.* **74**, 257–278 (1986).
- Stribley, D. P. & Read, D. J. *New Phytol.* **86**, 365–371 (1980).
- Bajwa, R. & Read, D. J. *New Phytol.* **101**, 459–467 (1985).
- Read, D. J. *Experientia* **47**, 376–391 (1991).
- Abuzinadah, R. A. & Read, D. J. *New Phytol.* **103**, 481–493 (1986).
- Abuzinadah, R. A. & Read, D. J. *New Phytol.* **112**, 61–68 (1989).
- Bliss, L. C. & Matveyeva, N. in *Arctic Ecosystems in a Changing Climate* (eds Chapin III, F. S., Jefferies, R. L., Reynolds, J. F., Shaver, G. R. & Svoboda, J.) 59–89 (Academic, San Diego, 1992).
- Haselwandter, K. & Read, D. J. *Oecologia* **53**, 352–354 (1982).

12. Chapin III, F. S., Fletcher, N., Kielland, K., Everett, K. R. & Linkins, A. E. *Ecology* **69**, 693–702 (1988).
13. Kinraide, T. B. *Plant Physiol.* **68**, 1327–1333 (1981).
14. Wyser, R. E. & Komar, E. *Plant Physiol.* **76**, 865–870 (1984).
15. Chapin III, F. S. & Bloom, A. J. *Oikos* **27**, 111–121 (1976).
16. Chapin III, F. S. & Tryon, P. R. *Holarct. Ecol.* **5**, 164–171 (1982).
17. Marion, G. M. & Kummerow, J. *Holarct. Ecol.* **13**, 50–55 (1990).
18. Koch, G. W., Bloom, A. J. & Chapin III, F. S. *Oecologia* **88**, 570–573 (1991).
19. Stewart, G. R., Lee, J. A. & Orebamjo, T. O. *New Phytol.* **72**, 539–546 (1973).
20. Nadelhoffer, K. G., Gilbin, A. E., Shaver, G. R. & Laundre, J. A. *Ecology* **72**, 242–253 (1991).
21. Haynes, R. J. & Goh, K. M. *Biol. Rev.* **53**, 465–510 (1978).
22. Chapin III, F. S., Bloom, A. J., Field, C. B. & Waring, R. H. *BioScience* **37**, 49–57 (1987).
23. Kroehler, C. J. & Linkins, A. E. *Oecologia* **85**, 424–428 (1991).
24. Chapin III, F. S. & Shaver, G. R. *Funct. Ecol.* **3**, 73–80 (1989).
25. Jones, B. N., Paabo, S. & Stein, S. J. *J. Liq. Chromatogr.* **4**, 565–586 (1981).
26. Epstein, E., Schmid, W. E. & Rains, D. W. *Plant Cell Physiol.* **4**, 79–84 (1963).
27. Wood, W. B., Wilson, J. H., Benbow, R. M. & Hood, L. E. *Biochemistry* 2nd edn (Benjamin/Cummings, Menlo Park, 1981).
28. Kedrowski, R. A. *J. Plant Nutr.* **6**, 989–1011 (1983).
29. Collins, C. H. & Lyne, P. M. *Microbiological Methods* 5th edn (Butterworths, London, 1984).
30. Bloom, A. J. & Chapin III, F. S. *Plant Physiol.* **68**, 1064–1067 (1981).

ACKNOWLEDGEMENTS. We thank M. Chapin, K. Dunkin, M. Firestone, J. Harte, S. Hobbie and G. Shaver for critical review of the manuscript. This research was supported by the NSF.

Disease resistance results from foreign phytoalexin expression in a novel plant

Rüdiger Hain*, Hans-Jörg Reif*, Elvira Krause*,
Ruth Langebartels*, Helmut Kindl†, Barbara Vornam†‡, Wilfried Wiese†,
Elmon Schmelzer§, Peter H. Schreier*, Ronald H. Stöcker* & Klaus Stenzel*

* Bayer AG, PF-E/Fu, Institut für Biotechnologie, D-5090 Leverkusen 1, Germany

† Biochemie, Fachbereich Chemie, Hans-Meerwein-Strasse, Philipps-Universität, D-3500 Marburg, Germany

‡ Max-Planck-Institut für Züchtungsforschung, Carl-von-Linné Weg, D-5000 Köln 30, Germany

ALTHOUGH phytoalexins^{1,2} have long been inferred to be important in the defence of plants against fungal infection^{1,2}, there are few reports showing that they provide resistance to infection. Several plants, including grapevine, synthesize the stilbene-type phytoalexin resveratrol^{3–7} when attacked by pathogens. Stilbenes with fungicidal potential are formed in several unrelated plant species, such as peanut (*Arachis hypogaea*), grapevine (*Vitis vinifera*) and pine (*Pinus sylvestris*)^{3,5,11–15}. Stilbene biosynthesis only specifically requires the presence of stilbene synthase^{6,9}. Furthermore, the precursor molecules for the formation of hydroxystilbenes are malonyl-CoA and *p*-coumaroyl-CoA, both present in plants⁹. To investigate the potential of stilbene biosynthetic genes in a strategy of engineering pathogen resistance, we isolated stilbene synthase genes from grapevine, where they are expressed at a high level, and transferred them into tobacco¹⁰. We report here that regenerated tobacco plants containing these genes are more resistant to infection by *Botrytis cinerea*. This is, to our knowledge, the first report of increased disease resistance in transgenic plants based on an additional foreign phytoalexin.

Complementary DNA⁸ was used as a probe to isolate a clone from a λEMBL4 genomic library of grapevine (*Vitis vinifera* var. Optima) containing two linked, full-length stilbene synthase genes (*Vst1* and *Vst2*, Fig. 1a). Direct gene transfer experiments into tobacco were done to prove the two genes isolated were stilbene synthase genes. The phage clone DNA was cotransferred with a plant selectable marker gene to protoplasts of tobacco (*Nicotiana tabacum* cv. Petit Havana SR1)¹⁰. Transgenic tobacco calluses and plants were regenerated^{9,10}. Southern blot analysis

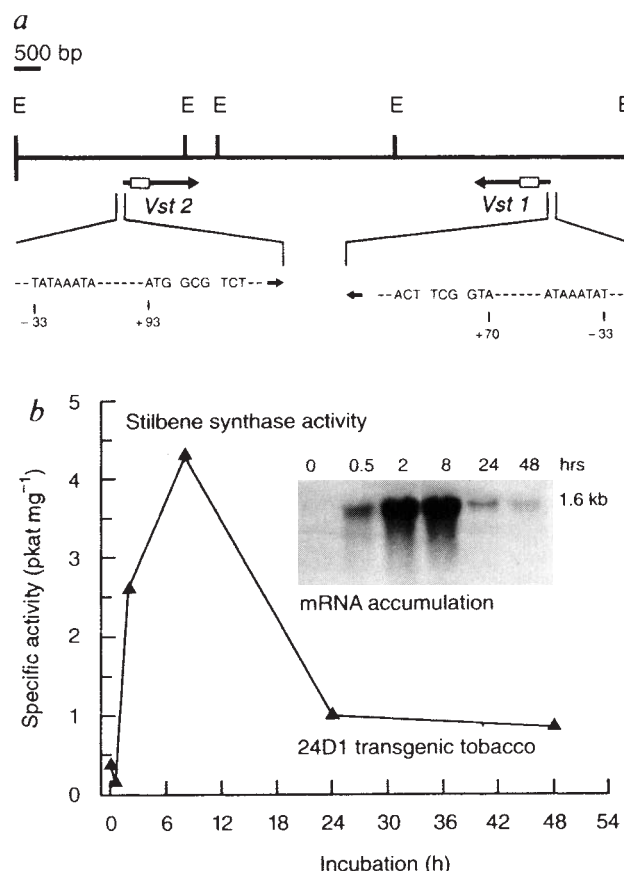


FIG. 1 a, Map of genomic DNA containing two stilbene synthase genes. Genomic DNA was isolated from partially purified nuclei¹⁶ of *Vitis vinifera* var. Optima. A library was constructed by size fractionation of *Nde*I cut genomic DNA and fragments of about 13 kilobases (kb) were cloned into the *Bam*HI site of λEMBL4 (ref. 17). Clones carrying stilbene synthase genomic sequences were identified by hybridization to a grapevine stilbene synthase cDNA⁶ and then restriction mapped. For each gene the TATA box at -33 and the length of the 5'-noncoding region on the mRNA is indicated. The box in the gene denotes the position of the intron. E = *Eco*RI. b, Profiles of stilbene synthase mRNA accumulation and enzymatic activities in transgenic tobacco suspension cultures following addition of elicitor.

METHODS. Suspension cultures were established by incubating transgenic tobacco callus material reduced to small pieces in liquid Linsmaier and Skoog medium²² (LS) containing 1 mg l⁻¹ NAA and 0.2 mg l⁻¹ kinetin on a rotary shaker at 100 r.p.m. over a period of three months. The medium was exchanged each week. For induction experiments, six flasks containing 50 ml LS medium each were inoculated with 5 ml of a 1-week-old suspension culture and after 3 days the cells were induced with 25 µg ml⁻¹ of a fungal elicitor preparation containing 11.25 µg glucose equivalents per ml from *Phytophthora megasperma* f. sp. glycinea (Pmg elicitor) known to induce a chimaeric peanut stilbene synthase gene^{7,9} in transgenic tobacco⁹. At each time point (0, 0.5, 2, 8, 24, 48 h) suspension cells were collected and crude protein extracts as well as RNA were prepared. Total RNA (20 µg) was northern blot analysed. As a probe, a stilbene synthase cDNA⁶ was used. Enzymatic activity was determined⁶ and the activity of 1 pkat was defined as incorporation of 1 pmol of *p*-coumaroyl-CoA into resveratrol at 30 °C.

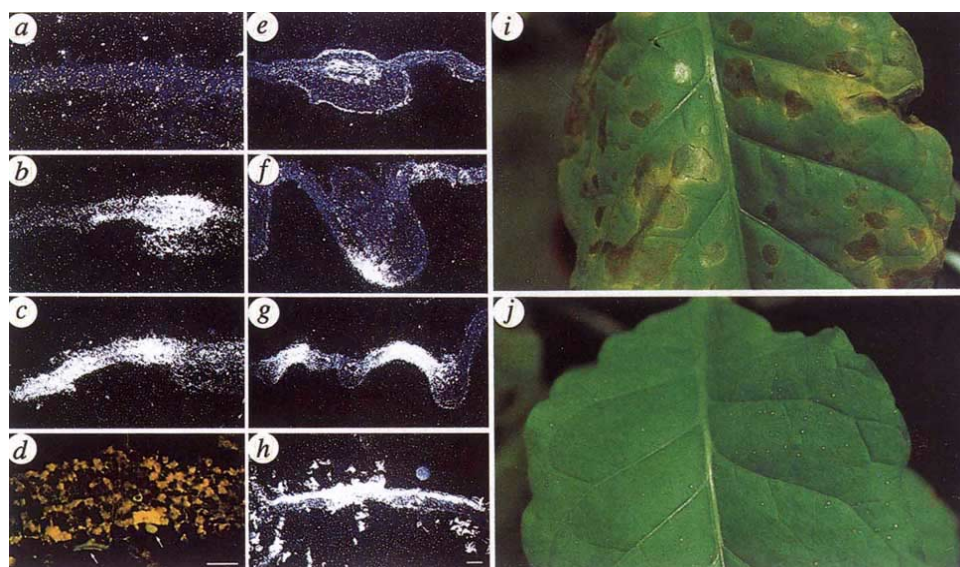
revealed one to two copies of the phage clone DNA to be integrated in the genome of tobacco (data not shown). Furthermore, *N. tabacum* cv. Petit Havana SR1 and *N. tabacum* var. Samsun were transformed with the stilbene synthase genes by *Agrobacterium tumefaciens*-mediated gene transfer¹⁰.

Transgenic tobacco calluses and plants were then analysed for regulated stilbene synthase gene expression. In grapevine, stilbene synthesis is inducible after fungal infection^{11,12}, elicitor treatment⁶, wounding or ultraviolet light irradiation^{12,14}. Assuming that similar regulation of grapevine stilbene synthase gene

‡ Present address: Institut für Forstgenetik und Forstpflanzenzüchtung Busgenweg 2, D-3400 Göttingen, Germany.

FIG. 2 *In situ* RNA hybridization experiments (a–h) and symptom density on tobacco SR1 and transgenic tobacco 23D17 (i, j). Localization of stilbene synthase mRNA by *in situ* hybridization in cross-sections of young leaflets of transgenic *Nicotiana tabacum* 24D1 (a, b, c), of *Vitis riparia* (e, f, g), and *Vitis vinifera* cv. optima plants (h) infected with the fungal pathogen *B. cinerea*. d, Micrograph taken under blue-epifluorescent light showing an infection site (yellow autofluorescence) and fungal structures (green fluorescence) stained with Blancophore in tissue of transgenic tobacco 24D1. Dark-field micrographs of *in situ* hybridizations done with stilbene synthase antisense probes show the gene expression in non-infected (a, e), and in infected tissues, 8 h (b, f) and 24 h (c, g, h) postinoculation. The arrows point at an infection site and fungal structures, respectively. Scale bars, 50 μ m.

METHODS. Small leaflets of shoot tips of sterile shoot cultures of transgenic *Nicotiana tabacum* 24D1, of *Vitis vinifera* cv. optima, and *Vitis riparia* were inoculated by dipping in a spore suspension (5×10^5 spores per ml) of *B. cinerea* and incubated in a moist chamber for 8 and 24 h. Fixation of the tissue, embedding in paraplast, and sectioning was as described earlier¹⁸. Serial cross-sections were hybridized *in situ* alternatively with sense and antisense RNA transcripts of stilbene synthase cDNA¹⁹. Transcripts were labelled with ³⁵S-labelled UTP. To identify



infection sites clearly, fungal structures in tissue sections were stained with the fluorescent dye Blancophore (Bayer) before hybridization. i, j, Tobacco plants were inoculated with *Botrytis* in the greenhouse. Photographs were taken 7 days after inoculation with *B. cinerea*. The leaves photographed are from trial B (see Table 1b, leaves number 4) where protection was 81.6% in disease reduction.

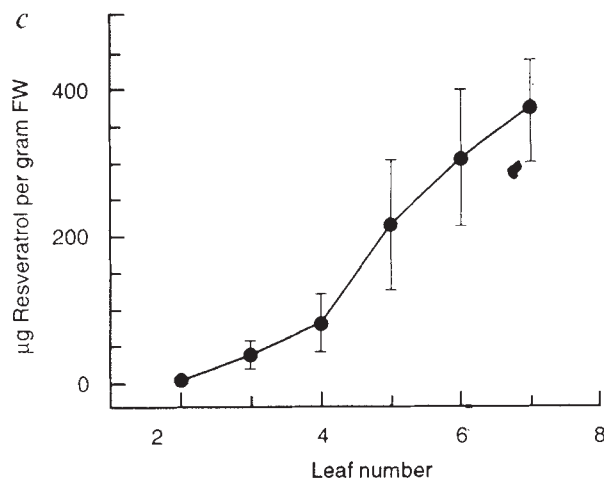
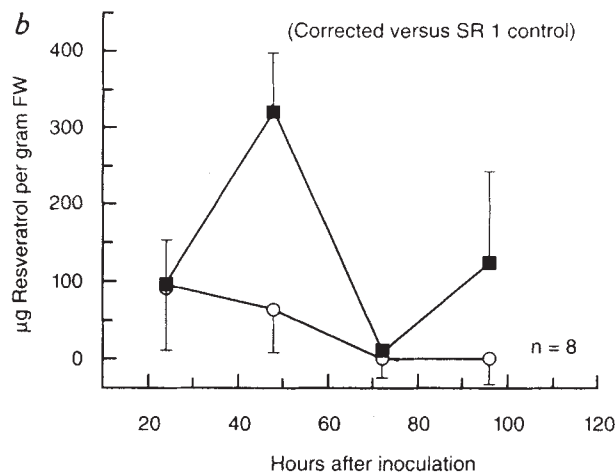
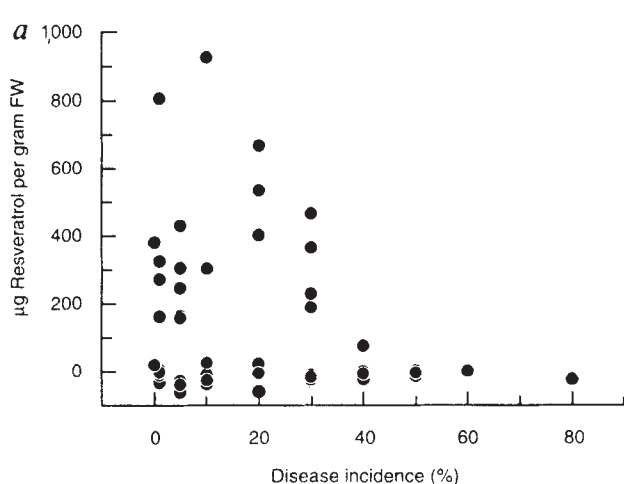


FIG. 3 a, Correlation of resveratrol production in leaves and susceptibility of transgenic tobacco to *B. cinerea* infection. Transgenic F_1 tobacco plants containing stilbene gene 1 (*Vst1*) and respective SR1 controls were inoculated with *B. cinerea* and after 24, 48, 72 and 96 h, half of the leaves were cut off and analysed for resveratrol content using the ELISA assay⁹. In the remaining half of the leaves the disease incidence was evaluated after 120 h. The amounts of resveratrol determined in the extracts of the leaves were plotted versus per cent disease incidence evaluated after 120 h. b, Comparison of the kinetics of resveratrol accumulation in leaves with high (○, >16% leaf area infected, mean 28%) and low (■, 0–15% leaf area infected, mean 11%) disease incidence. Leaves with high and low disease incidence after inoculation with *B. cinerea* were analysed for their resveratrol content⁹ during the incubation period. Leaves accumulating high levels of resveratrol 48 h after inoculation were more resistant to *B. cinerea* infection than controls of transgenic leaves which did not accumulate comparably high levels at that time. SR1 control mean, 23%, $n = 6$. c, Resveratrol content in leaves depending on leaf age. Older and younger leaves (numbers 2–7, respectively) of transgenic tobacco were analysed for their resveratrol content 48 h postinoculation (corrected versus SR1 control). Young leaves contain higher amounts of resveratrol than older leaves. Mean resveratrol content from two experiments with four leaves each.

TABLE 1 Influence of the stilbene synthase genes on the resistance of tobacco (*Nicotiana tabacum*) towards grey mould (*Botrytis cinerea*). Infection density of *Botrytis cinerea* on wild-type and transgenic tobacco plants

(a) Summary of biological testing in the greenhouse

Clone	Number of trials with difference in disease incidence* compared to controls			Average disease reduction (%)	Number of trials	Number of plants
	>25% increase	±25% no difference	>25% reduction			
Wild-type SR1 control						
15D15	2	3	1	-11.8† (±19.7)‡	6	86
23D17	0	1	5	49.0 (±10.2)	6	77
24D1	1	3	2	12.3 (±16.6)	6	78
Samsun control						
E1.3.1	0	0	3	41.4 (±7.2)	3	97

(b) Disease incidence on leaves of wild-type§ and transgenic tobacco with *B. cinerea* in dependence of leaf age

Clone	Diseased leaf area on leaves 1-6 (%)						Average diseased leaf area (%)	Disease reduction (%)	Number of plants
	1	2	old → young leaves		5	6			
			3	4					
Trial A									
SR1§	15.0	16.0	16.5	11.5	8.6	2.8	11.7	—	15
15D15	13.8	11.2	9.7	6.8	6.1	2.2	8.3	29.0	14
23D17	6.4	8.6	8.8	2.9	1.7	0.9	4.9	58.1	13
24D1	10.8	8.1	3.7	3.6	3.0	0.8	5.0	57.3	9
Trial B									
SR1§	11.3	13.6	9.5	13.9	3.8	0.8	8.7	—	21
15D15	16.6	16.8	9.7	5.1	2.6	0.4	8.5	2.3	21
23D17	3.7	4.0	1.1	0.7	0.4	0.0	1.6	81.6	14
24D1	7.3	9.9	8.8	8.3	6.6	0.8	6.7	23.0	14
Trial C									
Samsun§	11.4	15.2	19.7	18.8	9.0	1.8	12.6	—	32
E1.3.1	4.2	5.0	7.1	9.0	8.7	2.9	6.1	51.6	32

For the greenhouse tests, transgenic (15D15, 23D17, 24D1) and wild-type SR1 tobacco control plants were multiplied by *in vitro* shoot culture. Transgenic tobacco line 15D15, which served as transgenic control, contained a chimaeric peanut stilbene synthase gene⁸ and produced only low amounts of resveratrol. 15D15 showed no noticeable increased resistance to *B. cinerea*. Experiments with *N. tabacum* var. Samsun E1.3.1, generated by *A. tumefaciens* transformation, were done with F₁ generations. Transgenes and respective controls were incubated in tents and infected with spore suspensions of *B. cinerea* (10⁵ spores per ml). After incubation for 7 days at 19 °C and 100% relative humidity, infection density and symptom development were evaluated using six leaves (1-6, from bottom to top). For *Botrytis*, disease incidence is correlated with leaf age, that is, older leaves of wild-type tobacco plants are more susceptible than younger ones. Also in transgenic tobacco plants disease incidence was dependent on the leaf age. The difference in level of resistance between transgenic and wild-type tobacco was more pronounced on the younger leaves. This correlates with the highest amounts of resveratrol found in these leaves compared to old leaves (see Fig. 3c).

* Disease incidence = average of per cent diseased leaf area on leaves 1-6.

† Increase in disease incidence.

‡ s.e. of the mean in brackets.

§ Wild type.

expression takes place in tobacco, transgenic kanamycin-resistant tobacco calluses were treated with a fungal elicitor isolated from *Phytophthora megasperma* f. sp. glycinea. Kanamycin-resistant calluses accumulated stilbene synthase messenger RNA of the predicted size (data not shown), demonstrating that stilbene synthase genes can be induced in a heterologous plant system. Hence a signal pathway for such defence-related genes seems to be conserved between grapevine, peanut and tobacco. Quantitative northern blot analysis of RNA isolated from transgenic tobacco calluses revealed the presence of mRNA amounting to about 5% of the overall stilbene synthase mRNA accumulating in induced grapevine calluses.

To aid the analysis of stilbene synthase gene expression, we established suspension cultures of transgenic tobacco. Stilbene synthase mRNA accumulated transiently after induction with fungal elicitor (Fig. 1b). Enzymatic activity of stilbene synthase was assayed in parallel. In these cultures stilbene synthase activity followed the kinetics of stilbene synthase mRNA accumulation (Fig. 1b).

In leaves of transgenic tobacco inoculated with *B. cinerea* a transient stilbene synthase mRNA accumulation was also found (data not shown). *In situ* RNA hybridization demonstrated a specific expression of the genes in the vicinity of infection sites to high local levels in both grapevine and transgenic tobacco (Fig. 2a-h).

To assess the effect of pathogen-inducible stilbene synthase gene expression on disease resistance, tobacco plants were inoculated with *B. cinerea*. This fungus was chosen as pathogen because, in certain varieties of grapevine (*Vitis vinifera*), a positive correlation was found between the concentration of resveratrol formed when attacked by the pathogen and the resistance against the fungus¹¹.

Four transgenic tobacco lines were tested for enhanced disease resistance to *B. cinerea* in the greenhouse (Table 1a, b; Fig. 2i, j). Two transgenic lines (shoot cultures of line 23D17 and R₁ generations of E1.3.1) were reproducibly more resistant to *B. cinerea* than the wild-type and transgenic 15D15 controls. Transgenic line 24D1 was not reproducibly more resistant than control tobacco. Analysis of the kinetics of stilbene synthase mRNA accumulation in leaves of both lines (24D1 and 23D17) revealed that line 23D17 accumulated high amounts of specific mRNA more rapidly. In 23D17 shortly after inoculation stilbene synthase mRNA was detectable, reaching a maximum between 6 and 24 h (data not shown). In line 24D1 there was no stilbene synthase mRNA detectable a short time after inoculation and maximum accumulation was observed between 24 and 48 h postinoculation. But at maximum stilbene synthase mRNA accumulation the levels of specific mRNA were comparable in 24D1 and 23D17 (data not shown). The reason for the difference in the speed of stilbene synthase gene activation in these lines

is not known, but might be due to position effects resulting from different locations of the transferred DNA in the genome of tobacco.

The level of resistance of plants mediated by fungitoxic substances such as the phytoalexin resveratrol is believed to depend on the ability to produce a high concentration of the compound within a short time after infection^{1,2,9}. For this reason we analysed the levels of resveratrol and the kinetics of resveratrol accumulation in leaves of transgenic tobacco after fungal infection. The results support the notion that a resveratrol-based increased resistance depends on rapid synthesis of high amounts of the phytoalexin in transgenic plants (Fig. 3a–c). In transgenic SR1 plants containing stilbene synthase genes from grapevine up to 400 µg resveratrol per g fresh weight were detected (Fig. 3a–c). This concentration of resveratrol also affected fungal growth *in vitro* (data not shown). A correlation between amounts of resveratrol and disease incidence in transgenic tobacco leaves could be demonstrated (Fig. 3a). Furthermore a comparison of transgenic SR1 tobacco leaves with high or low disease incidence with the accumulation of resveratrol within the leaves indicates that a high concentration of resveratrol 48 h after inoculation is required for enhanced disease resistance (Fig. 3b).

Phytoalexins are vitally important in certain pathogen resistance interactions; in particular, virulence of *Nectria haematococca* on pea depends on its ability to degrade the host phytoalexin²⁰. Furthermore, an avirulent *Cochliobolus heterostrophus* transformed with a gene encoding the phytoalexin-degrading enzyme of *Nectria* became virulent on pea²¹. Thus our work confirms and extends the findings that phytoalexins can determine resistance by moving a phytoalexin from one species to another rather than moving detoxification genes from one pathogen to another. This demonstrates an increased disease resistance of transgenic plants expressing defence-related stilbene synthase genes and raises the possibility of induction of resistance against fungal pathogens by inserting foreign genes coding for phytoalexins. Moreover, we provide the first evidence, to our knowledge, that biosynthetic pathways of plants can be successfully modified or altered by recombinant DNA technology, resulting in pathogen-induced synthesis of a foreign phytoalexin in tobacco plants. Because phytoalexins are fungitoxic substances acting against a variety of different pathogenic fungi, it seems reasonable to assume that there is a general relevance of the described system for other host-pathogen interactions. □

Received 1 September; accepted 16 November 1992.

- Ingham, J. *Bot. Rev.* **38**, 343–424 (1972).
- Ebel, J. A. *Rev. Phytopath.* **24**, 235–264 (1986).
- Gorham, J. *Prog. Phytochem.* **6**, 203–252 (1980).
- Ingham, J. *Phytochemistry* **15**, 1791–1793 (1976).
- Pont, V. & Pezet, R. *J. Phytopath.* **130**, 1–8 (1990).
- Schöppner, A. & Kindl, H. *FEBS Lett.* **108**, 345–352 (1979).
- Schröder, G., Brown, J. W. S. & Schröder, J. *Eur. J. Biochem.* **172**, 161–169 (1988).
- Melchior, F. & Kindl, H. *Arch. Biochem. Biophys.* **288**, 552–557 (1991).
- Hain, R., Bieseler, B., Kindl, H., Schröder, G. & Stöcker, R. *Pl. molec. Biol.* **15**, 325–336 (1990).
- Hain, R. *et al. Molec. Gen. Genet.* **199**, 161–168 (1985).
- Stein, U. & Blaich, R. *Vitis* **24**, 75–87 (1985).
- Langcake, P. *Physiol. Pl. Path.* **9**, 77–86 (1976).
- Langcake, P. & Pryce, R. J. *Experientia* **33**, 151–152 (1977).

- Langcake, P. & Pryce, R. J. *Phytochemistry* **16**, 1193–1196 (1977).
- Dercks, W. & Creasy, L. L. *Physiol. molec. Pl. Path.* **34**, 189–202 (1989).
- Bedbrook, J. *Pl. molec. Biol. Newsl.* **2**, 24 (1981).
- Frischauf, A. M., Lebrach, H., Poustka, A. & Murray, N. J. *molec. Biol.* **170**, 827–842 (1983).
- Cuyper, B., Schmelzer, E. & Hahlbrock, K. *Molec. Pl.-Microbe Interact.* **1**, 157–160 (1988).
- Schmelzer, E., Krüger-Lebus, S. & Hahlbrock, K. *Pl. Cell* **1**, 993–1001 (1989).
- Tegtmeier, K. J., VanEtten, H. D. *Phytopathology* **72**, 608–612 (1982).
- Schäfer, W., Straney, E., Ciuffetti, L., VanEtten, H. D. & Yoder, O. C. *Science* **246**, 247–249 (1989).
- Linsmaier, E. M. & Skoog, F. *Physiol. Pl.* **18**, 100–127 (1965).

ACKNOWLEDGEMENTS. We thank R. Walden for his advice and critical reading of the manuscript, W. Zitzmann for fungal elicitor and M. Gilles, E. Oehmen, A. Koopmann and I. Cicerello for technical assistance.

Ocular regression conceals adaptive progression of the visual system in a blind subterranean mammal

Howard M. Cooper*, Marc Herbin*† & Eviatar Nevo‡

* Cerveau et Vision, INSERM Unité 371, 69500 Bron, France

† Laboratoire d'Anatomie Comparée, Muséum National d'Histoire Naturelle, 75005 Paris, France

‡ Institute of Evolution, University of Haifa, Haifa, Israel

THE mole rat, *Spalax ehrenbergi*, is an extreme example of natural visual degeneration in mammals: visual pathways are regressed and incomplete¹, and the absence of visual cortical potentials or an overt behavioural response to light have led to the conclusion that *Spalax* is completely blind^{2–4}. But structural and molecular investigations of the atrophied, subcutaneous eye suggest a functional role for the retina in light perception^{5,6}, and entrainment of circadian locomotor and thermoregulatory rhythms by ambient light demonstrates a capacity for photoperiodic detection^{7–9}. We report here that severe regression of thalamic and tectal structures involved in form and motion perception is coupled to a selective hypertrophy of structures subserving photoperiodic functions. As an alternative to the prevalent view that ocular regression results from negative or nonselective evolutionary processes^{10–12}, the differential reduction and expansion of visual structures in *Spalax* can be explained as an adaptive response to the underground environment.

We used retrograde and anterograde tracing techniques to study retinal ganglion cells and retinal projections in the blind

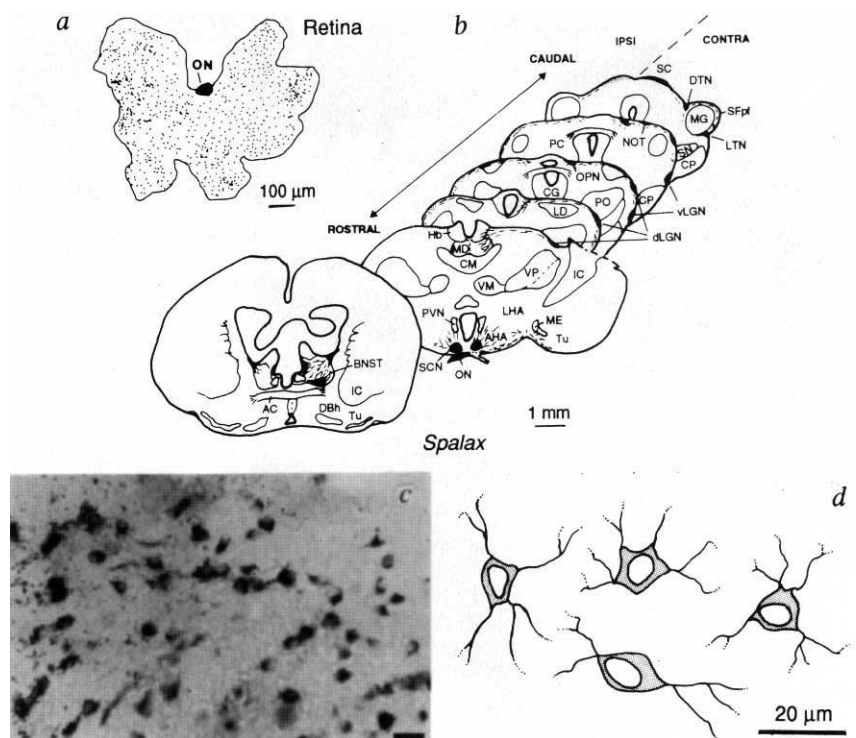
mole rat, which has the most rudimentary eye of any mammal^{5,13}. We found less than 900 ganglion cells, which is inferior to the number of sensory fibres of the third eye (pineal complex or frontal organ) in certain vertebrates¹⁴. Ganglion cells are uniformly distributed throughout the retina and form a morphologically homogeneous population (Fig. 1). This lack of retinal heterogeneity could represent a global effect on the morphology of the entire ganglion cell population or may reflect the selective elimination of particular classes of cells.

Intraocular injection of anterograde tracer in the minute eye shows that bilateral retinal projections to all visual nuclei are present in *Spalax*, although compared with non-fossorial rodents, visual pathways are not simply scaled down as a function of eye size. Retinal input to the dorsal lateral geniculate nucleus (dLGN; Figs 1 and 2) is restricted to a thin sheet of label on the superficial dorsal thalamus, within a nucleus no more than 3–5 cells wide. As in other mammals, these cells project to the primary visual cortex¹⁵. The pretectum (PRT), ventral lateral geniculate nucleus (vLGN) and accessory optic system (AOS) also contain a sparse retinal input (Fig. 1). Despite the substantial reduction in absolute size of these structures in *Spalax* as compared with other rodents (87–93%), relative retinal input is proportionately equivalent in both aboveground and underground species (Fig. 3b–c).

The superior colliculus in *Spalax* receives a sparse, superficial retinal innervation and is severely reduced in both absolute and relative size (Figs 1 and 2). The superficial lamina are collapsed into a single layer 1–2 cells thick. In other rodents, the superior colliculus typically receives the densest retinal innervation (up to 74%), whereas in *Spalax*, the superior colliculus accounts for only 20% of the total retinal projection, a relative decrease of 50%. The absolute reduction in size is even more noticeable: the volume of the retino-recipient layers in *Spalax* is only 3%

FIG. 1 *a*, After horseradish peroxidase (HRP) application to the exposed optic nerve, a total of 823 retrogradely labelled ganglion cells were found homogeneously distributed throughout the retina (surface area, 1.3 mm²; ON, optic nerve head). *c*, High-power photomicrograph of HRP-labelled ganglion cells (scale bar 20 μ m), shown in greater detail in the camera lucida drawings in *d*. The cell soma size distribution is unimodal, with diameters ranging from 6–15 μ m (mean, 10.37, s.d. = 1.79). Although this method may not have fully revealed cell morphology, ganglion cells are all similar in shape, with a large nucleus, scant cytoplasm and limited dendritic arborization. We also observe that all optic nerve fibres are unmyelinated (personal observation). *b*, After intraocular injection of cholera toxin subunit B conjugated with HRP (CT-HRP); more than 20 different structures in the hypothalamus, thalamus, mesencephalon and basal telencephalon receive a projection from the retina (labelled fibres, fine lines; terminal-like label, shaded areas). The encapsulated region of the bed nucleus of the stria terminalis (BNST) contains a diffuse plexus of labelled fibres and terminals, which also extends to several adjacent regions (anterodorsal, anteroventral, thalamic paraventricular and mediodorsal nuclei, lateral habenula). The suprachiasmatic nucleus (SCN) receives a dense bilateral innervation, and retinal fibres are present in the anterior hypothalamic area (AHA), lateral hypothalamic area, subparaventricular zone and retrochiasmatic area. A few labelled fibres are observed in the basal telencephalon near the amygdaloid region and olfactory tubercle³¹. The vLGN and dLGN are extremely thin, superficial structures and no cytoarchitectural subdivisions could be distinguished. In the pretectum, the olivary pretectal nucleus and the nucleus of the optic tract receive retinal projections. The superior colliculus is highly reduced. A sparse retinal projection to the dorsal (DTN) and lateral terminal nuclei (LTN) of the accessory optic system is also present. The density of label in ipsilateral thalamic and tectal nuclei is less than 2% of the contralateral label, in contrast to the SCN which receives a balanced bilateral innervation.

METHODS. Six mole rats from the Anza region of central Israel were studied. This population belongs to the recently evolved chromosomal species ($2n = 60$) of the *Spalax ehrenbergi* superspecies (Spalacidae, Rodentia) which actively speciated during the Pleistocene into four sibling species. Animals were anaesthetized (30–40 mg kg⁻¹, ketamine hydrochloride; 2 mg kg⁻¹, xylazine) before receiving an intraocular injection of a 0.2% solution of CT-HRP dissolved in 0.5–1.5 μ l distilled water. An incision was made in the skin slightly above the location of the subcutaneous eye and the HRP conjugate injected with a 50 μ m tipped glass pipette sealed to the needle of a Hamilton syringe. Care was taken to avoid damage to the retina which fills most of the eyecup. After survival periods of 48 to 72 h, animals received a lethal dose of anaesthetic (sodium pentobarbitol) and were subsequently perfused through the heart with 300 ml warm saline (0.9%)



followed by 1 l cold 4% paraformaldehyde in 0.1 M phosphate buffer and a post-fixation rinse of 10% sucrose in the same buffer. Coronal sections (30–40 μ m) were cut on a freezing microtome and reacted for HRP³². To obtain retrograde labelling of retinal ganglion cells, the intact eye was removed during anaesthesia in some animals and a crystal of HRP placed on the exposed optic nerve. After 2 min the HRP was washed away and the eye allowed to incubate in oxygenated Ringer's solution. The retina was then removed, reacted as above for HRP and flat mounted. **Abbreviations:** AC, anterior commissure; AHA, anterior hypothalamic area; BNST, bed nucleus of the stria terminalis; CG, central periaqueductal grey; CM, centromedian nucleus; CP, cerebral peduncle; DBH, diagonal band of Broca, horizontal limb; dLGN, dorsal lateral geniculate nucleus; DTN, dorsal terminal nucleus; Hb, habenula; IC, internal capsule; LD, laterodorsal nucleus; LHA, lateral hypothalamic area; LTN, lateral terminal nucleus; MD, mediodorsal nucleus; ME, medial amygdala; MG, medial geniculate nucleus; NOT, nucleus of the optic tract; OPN, olivary pretectal nucleus; ON, optic nerve; PC, posterior commissure; PO, posterior complex of the thalamus; PVN, paraventricular hypothalamic nucleus; SC, superior colliculus; SF-pf, superior fasciculus, posterior fiber branch; SN, substantia nigra; Tu, olfactory tubercle; vLGN, ventral lateral geniculate nucleus; VM, ventromedial nucleus; VP, ventroposterior nucleus.

of that observed in the hamster and other rodents (Fig. 3*a–c*).

In contrast to the severe hypoplasia, poorly differentiated cytoarchitecture and extremely reduced retinal input of thalamic and mesencephalic structures, the morphology and retinal innervation of the suprachiasmatic nucleus of the hypothalamus (SCN; Figs 1 and 2) in *Spalax* is identical to that in other rodents^{16,17}. The relative size of the retinal projection to the SCN is, however, significantly expanded. The density of retinal input to the SCN in *Spalax* accounts for nearly 20% of the total visual projection, which is equivalent to the amount of label found in the superior colliculus and exceeds that of any other primary visual structure (Fig. 3). In other species, the density of retinal projections to the SCN constitutes a minute fraction (less than 1%) of the total projection. Retrograde tracing studies in rodents have shown that only 30–130 ganglion cells from each retina project to the SCN, compared with the total ganglion cell population of 65,000 (mouse)¹⁸ to 100,000 (hamster)¹⁹. This

suggests that in *Spalax*, ganglion cells projecting to the hypothalamus constitute a substantial proportion of the total ganglion cell population.

Although overlooked in previous studies of subterranean mammals^{1,20,21}, an exceptionally dense retinal projection is observed in the bed nucleus of the stria terminalis (BNST). Contralaterally, labelled retinal fibres profusely invade several regions within the encapsulated part of the BNST, although some fibres continue from this region to the anterodorsal and anteroventral nuclei, the mediodorsal nucleus, the lateral habenula, and the thalamic paraventricular nucleus (Figs 1 and 2). Compared to other rodents in which these regions merely contain a few scant retinal fibres²², in the mole rat these pathways are hypertrophied and show both a relative and absolute expansion (Fig. 3).

Despite an acute peripheral reduction these results confirm the conservation of a basic pattern of primary visual

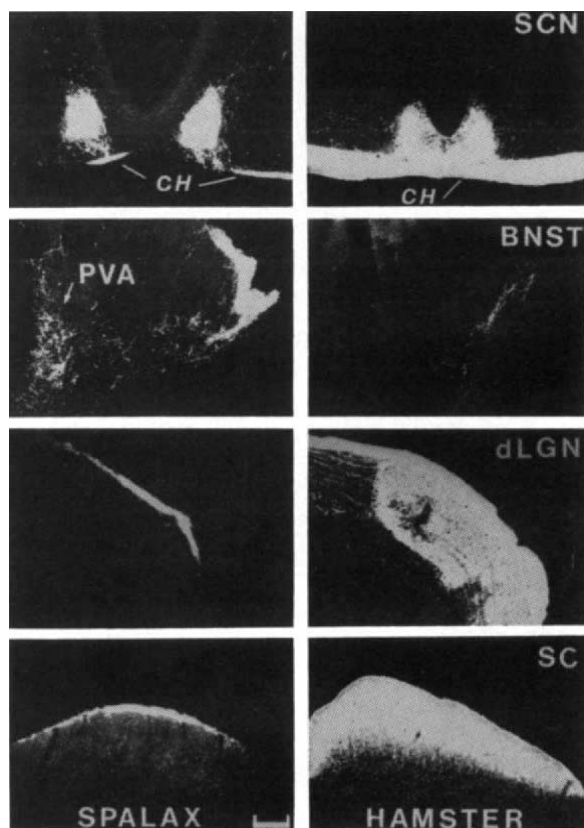


FIG. 2 Dark-field illustrations of retinal projections to contralateral visual structures in the blind mole rat (left side) and in the hamster (right side) following intraocular injection of CT-HRP. For direct comparison, all photographs are scaled to the same size (scale bar, 300 μ m). The size of the suprachiasmatic nucleus (SCN) is similar in the mole rat and in the hamster. In both species, anterogradely labelled HRP terminals fill the entire nucleus, although the density is greatest in the ventral region. Note the difference, however, in the size of the optic chiasm (CH). The retinal projection to the lateral and medial regions of the bed nucleus of the stria terminalis (BNST) is hypertrophied in *Spalax* as compared to the presence of a few scant fibres within this region in the hamster. In *Spalax*, labelled fibres also invade the paraventricular thalamic nucleus (PVA) along the wall of the third ventricle. The extent of the dorsal lateral geniculate nucleus (dLGN) is drastically reduced in *Spalax* and the width of the entire nucleus is no greater than the width of the optic tract in the hamster (80–100 μ m wide). The superior colliculus (SC) in *Spalax* is also severely degenerate and the superficial lamina (stratum zonale, stratum griseum and stratum opticum) are fused into a single layer 50 μ m thick. The total volumes (in mm^3) occupied by retinal fibres within of the different visual structures in *Spalax* are: SCN, 0.0280; BNST, 0.0411; dLGN, 0.0388, SC, 0.0514; and in the hamster: SCN, 0.0253; BNST, 0.0012; dLGN, 0.5884; SC, 1.9840.

organization, and demonstrate that evolutionary regression is uniquely focused on non-hypothalamic structures involved in the analysis of spatial parameters and visuomotor integration. This selective reduction in *Spalax* is coherent with the inability of the subcutaneous eye and degenerate lens to focus an image on the retina, and with the lack of extraocular eye muscles^{5,13}. By contrast, the diffuse, low-level illumination of the subcutaneous eye is entirely compatible with the function of ambient light detection by the SCN. In mammals, photic input to the SCN synchronizes hormonal, physiological and behaviour rhythms to the photoperiodic light cycle²³. SCN neurons respond to large-field, low-level, diffuse light (0.1–500 lux)²³ and a single, short daily exposure (<3 s) is sufficient to synchronize and maintain the complete 24-h cycle of circadian locomotor activity²⁴. Photic entrainment of locomotor and thermoregulatory activity constitutes the exclusive visual capacity in

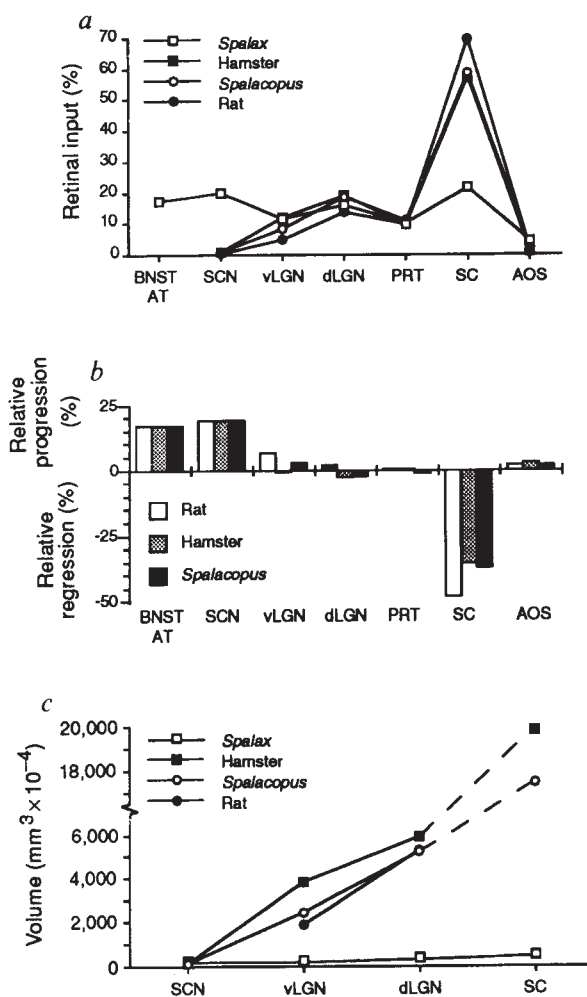


FIG. 3 *a*, Relative degree of retinal input to primary visual structures compared to the total quantity of retinal projections in *Spalax*, hamster, rat and *Spalacopus cyanus* (South American Octodontidae, 'cururo'). All these rodents are of similar body size (120–140 g). The density of retinal projections in each primary visual structure was quantified using a previously described method¹⁷. *b*, Relative degree of change in the proportions of retinal input to different primary visual structures in *Spalax* compared to measures obtained in the other rodents. A relative progressive development in *Spalax* is seen in structures involved in photoperiodic functions (SCN, BNST). The main regressive feature is the drastic reduction of retinal input to the superior colliculus. The relative size of other visual structures in *Spalax* are unmodified compared to that of the other species. *c*, Comparison of the absolute size (volume, $\text{mm}^3 \times 10^{-4}$) of visual structures in *Spalax* and other rodents. The size of the SCN is equivalent in all species. The vLGN and dLGN are reduced by 87–93% in *Spalax*. The retino-recipient layers of the superior colliculus are reduced by 97%. Values for the rat are adapted from Toga and Collins³³ and Sugita and Otani³⁴.

Spalax^{2,7–9,25} and we have previously shown this response is abolished if the eyes are removed^{7,8}. Because a considerable amount of light penetrates mammalian skin, activity at the ground's surface while excavating dirt through mounds could provide adequate light stimulation to the retina. For solitary species like *Spalax*, in which individuals are spatially isolated by exclusive territorial domains, photoperiodic synchronization of neuroendocrine physiology and behaviour may be a necessary prerequisite for successful reproduction²⁵.

The BNST is also involved in neuroendocrine regulation and reproductive functions²⁶ and plays a crucial role in seasonal temperature regulation. Vasopressinergic innervation of the lateral septum by the BNST controls both the maintenance of normal body temperatures and hypothermia^{27–29}. In *Spalax*, the

thermoregulatory capacity to adapt to cold temperatures increases with the decrease in day length preceding winter conditions² and retinal innervation of the BNST provides a putative pathway for mediation of this response. The presence of melatonin receptors in the SCN, BNST and several rostral thalamic regions³⁰, which in *Spalax* also contain retinal fibres (anteroventral, thalamic paraventricular and lateral habenular nucleus), provides further evidence for the involvement of these structures in photoperiodic functions.

We suggest that the photoperiodic system, sustaining appropriate reproductive and thermoregulatory responses, has been selectively expanded, whereas the acute metabolic burden of maintaining a large eye and non-functional 'image forming' visual system provides the underlying evolutionary impetus for their morphological regression. Thus, ocular regression, linked to selective progression of photoperiodic pathways, represents one aspect of the unique range of adaptations²⁵ that subterranean mammals have evolved to cope with environmental constraints imposed by the underground niche. □

Received 3 August; accepted 9 November 1992.

1. Bronchti, G., Rado, R., Terkel, J. & Wollberg, Z. *Devl Brain Res.* **58**, 159-170 (1991).
2. Haim, A., Heth, G., Pratt, H. & Nevo, E. *J. exp. Biol.* **107**, 59-64 (1983).
3. Heit, P., Bronchti, G., Wollberg, Z. & Scheich, H. *Neuroreport* **2**, 735-738 (1991).
4. Necker, R., Rehkämper, G. & Nevo, E. *Neuroreport* **3**, 505-508 (1992).
5. Sanyal, S., Jansen, H. G., De Grip, W. J., Nevo, E. & De Jong, W. W. *Invest. Ophthalmol. vis. Sci.* **31**, 1398-1404 (1990).

6. de Jong, W. W., Hendriks, W., Sanyal, S. & Nevo, E. in *Evolution of Subterranean Mammals at the Organismal and Molecular Levels* (eds Nevo, E. & Reig, O. A.) 383-395 (Liss, New York, 1990).
7. Pevet, P., Heth, G., Haim, A. & Nevo, E. *J. exp. Zool.* **232**, 41-50 (1984).
8. Heth, G., Pevet, P., Nevo, E. & Beiles, A. *J. exp. Biol.* **238**, 1-9 (1986).
9. Rado, R., Gev, H. & Terkel, J. *Israel J. Zool.* **35**, 105-106 (1988).
10. Wright, S. *Am. Nat.* **98**, 65-69 (1964).
11. Brace, C. L. *Am. Nat.* **97**, 39-49 (1963).
12. Wilkens, H. *Evolution* **25**, 530-544 (1971).
13. Cei, G. *Monitore zool. ital.* **55**, 69-88 (1946).
14. Collin, J.-P. & Oksche, A. in *The Pineal Gland* (ed. Reiter, R. J.) 27-67 (CRC Press, Boca Raton, 1981).
15. Cooper, H. M., Herbin, M. & Nevo, E. *J. comp. Neurol.* (in the press).
16. Cassone, V. M., Speh, J. C., Card, J. P. & Moore, R. Y. *J. biol. Rhythms* **3**, 71-91 (1988).
17. Magnin, M., Cooper, H. M. & Mick, G. *Brain Res.* **488**, 390-397 (1989).
18. Balkema, G. W. & Dräger, U. C. *Vis. Neurosci.* **4**, 593-604 (1990).
19. Pickard, G. E. *J. comp. Neurol.* **211**, 65-83 (1982).
20. Kudo, M., Nakamura, Y., Morizumi, T., Tokumo, H. & Kitao, Y. *Neurosci. Lett.* **93**, 176-180 (1988).
21. Lund, R. D. & Lund, J. S. *Expl Neurol.* **13**, 302-316 (1965).
22. Levine, J. D., Weiss, M. L., Rosenwasser, A. M. & Miselis, R. R. *J. comp. Neurol.* **306**, 344-360 (1991).
23. Meijer, J. H. & Reitveld, W. J. *Physiol. Rev.* **69**, 671-707 (1989).
24. Nelson, D. E. & Takahashi, J. S. *J. Physiol.* **439**, 115-145 (1991).
25. Nevo, E. in *Evolutionary Biology* Vol. 25 (eds Hecht, M. K., Wallace, B. & MacIntyre, R. J.) 1-125 (Plenum, New York, 1991).
26. Beltramo, C. & Taleisnik, S. *Neuroendocrinology* **30**, 238-242 (1980).
27. Cooper, K. E., Kasting, N. W., Lederis, K. & Veale, W. L. *J. Physiol., Lond.* **295**, 33-45 (1979).
28. Kasting, N. M. & Martin, J. D. *Brain Res.* **258**, 127-132 (1983).
29. Pévet, P., Buijs, R. M., Masson-Pévet, M. & Canguilhem, B. in *Fundamentals and Clinics in Pineal Research* (eds Trentini, G. P., de Gaetani, C. & Pévet, P.) 221-235 (Raven, New York, 1987).
30. Weaver, D. R., Rivkees, S. A., & Reppert, S. M. *J. Neurosci.* **9**, 2581-2590 (1989).
31. Cooper, H. M., Mick, G. & Magnin, M. *Brain Res.* **477**, 350-357 (1989).
32. Cooper, A. M. & Cowey, A. *Neuroscience* **35**, 335-344 (1990).
33. Toga, A. W. & Collins, R. C. *J. comp. Neurol.* **199**, 443-464 (1981).
34. Sugita, S. & Otani, K. *Expl. Neurol.* **82**, 413-423 (1983).

ACKNOWLEDGEMENTS. We thank S. Richard, N. Boyer, J.-L. Borach, P. Giroud, C. Urquizar and C. Bullier for technical assistance. This work was supported by grants to H.M.C. from the Rhone-Alpes Neuroscience Program, INSERM and the CNRS, to M.H. from the Federation des Aveugles, and the Museum d'Histoire Naturelle, Paris, BQR '92, and to E. Nevo from the 'Israeli Discount Bank Chair of Evolutionary Biology' and the Ancell-Teicher Research Foundation for Genetics and Molecular Evolution established by Florence and Theodore Baumritter of New York.

Pharmacology of GABA receptor Cl^- channels in rat retinal bipolar cells

Andreas Feigenspan, Heinz Wässle & Joachim Bormann*

Max-Planck-Institut für Hirnforschung, Deutschordenstrasse 46, W-6000 Frankfurt am Main 71, Germany

γ -AMINOBUTYRIC acid (GABA), a major inhibitory neurotransmitter in the mammalian nervous system, is known to operate bicuculline-sensitive Cl^- channels through GABA_A receptors and bicuculline-insensitive cation channels through GABA_B receptors^{1,2}. Recent observations indicate that the retina may contain GABA receptors with unusual pharmacological properties³⁻⁵. Here we report that GABA gates bicuculline-insensitive Cl^- channels in rod bipolar cells of the rat retina, which were not modulated by flunitrazepam, pentobarbital and alphaxalone and were only slightly blocked by picrotoxinin. Moreover, the GABA_B receptor agonist baclofen, and the antagonist 2-hydroxysaclofen had no effect. The underlying single-channel conductance was 7 pS and the open time 150 ms. These values are clearly different from those obtained for GABA_A receptor channels recorded in other neurons of the same preparation, and in other parts of the brain^{1,6-8}. The bicuculline- and baclofen-insensitive GABA receptors were activated selectively by the GABA analogue *cis*-4-aminocrotonic acid (CACA)⁹. Hence they may be similar to those receptors termed GABA_C receptors¹⁰.

Patch-clamp experiments were performed on organotypic slice cultures¹¹ prepared from 6-day-old rat retinae¹². Neurons were identified by injecting Lucifer Yellow using the patch pipette¹². Small bipolar cells and larger multipolar cells, presumably amacrine cells, were routinely obtained. All cells displayed inward currents in response to externally applied GABA

TABLE 1 Pharmacology of retinal GABA receptors

Drugs		Bicuculline-insensitive	Bicuculline-sensitive
Flunitrazepam	(1 μM)	-2 ± 10 (6)	120 ± 73 (9)
Pentobarbital	(50 μM)	5 ± 13 (5)	247 ± 135 (4)
Picrotoxinin	(10 μM)	-2 ± 15 (11)	-52 ± 10 (9)
	(100 μM)	-19 ± 24 (8)	-96 ± 4 (3)
Alphaxalone	(1 μM)	-8 ± 13 (5)	62 ± 66 (10)
Strychnine	(5 μM)	$+1 \pm 18$ (4)	-52 ± 7 (4)

Numbers are the percentage changes of GABA-induced whole-cell currents observed upon including various drugs in the GABA (20 μM) solution. Bicuculline-insensitive currents were recorded from bipolar cells in the presence of 100 μM bicuculline. The open-channel blocker picrotoxinin was applied twice (interval, 2 min) and the second application was evaluated. Numbers are given as means $\pm$ s.d. (*n* cells).

(20 μM) when voltage-clamped at negative membrane potentials. In most bipolar cells, high concentrations of bicuculline (100 μM) were unable to completely abolish the GABA response (Fig. 1a). The bicuculline-resistant component, when present, was from 17% to 89% (mean $52\% \pm 23\%$; 28 cells) of the original GABA-induced current. In contrast, GABA receptors in multipolar cells were exquisitely sensitive to bicuculline (Fig. 1b), the half-maximal inhibitory concentration being only 0.4 μM (not shown). Thus retinal bipolar cells contain GABA receptors which are different from the known GABA_A subtypes. As rodent bipolar cells are mainly targeted by rod photoreceptors¹³, the cells studied here were probably rod bipolar cells.

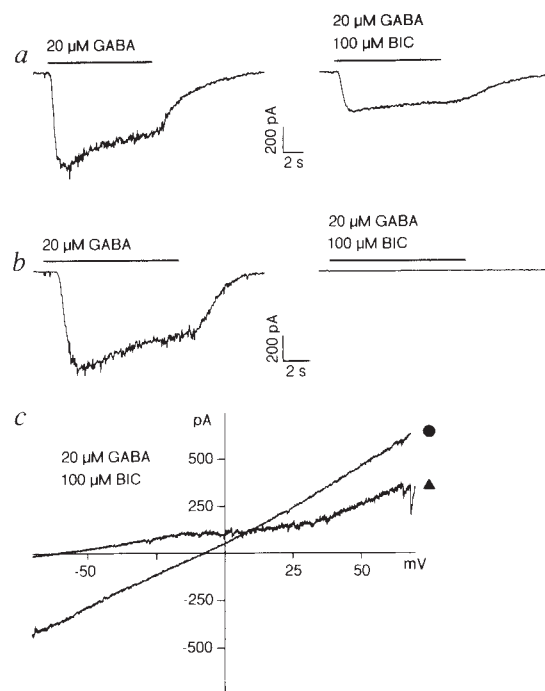
GABA-induced whole-cell currents recorded with symmetrical Cl^- in the presence of bicuculline revealed reversal potentials around 0 mV (mean $-4 \text{ mV} \pm 3 \text{ mV}$; 4 cells). Upon partial substitution of the intracellular Cl^- by equimolar amounts of gluconate, the reversal potential was shifted to more negative potentials, as expected for Cl^- -selective channels (Fig. 1c).

To characterize the pharmacology of bicuculline-insensitive GABA receptors further, we tested the effects of flunitrazepam,

* To whom correspondence should be addressed.

FIG. 1 Isolation of bicuculline-insensitive Cl^- currents in a subpopulation of retinal neurons. Cells were voltage-clamped at -70 mV and filled with Lucifer Yellow contained in the patch pipette¹². Bipolar cells having the characteristic morphology were usually very small (capacitance 3.3–6.1 pF) and showed no voltage-gated Na^+ currents. Multipolar cells, presumably amacrine cells, were larger (9.2–14.7 pF) and most of them (60%) expressed the Na^+ currents¹². *a*, A bipolar cell in which the GABA-evoked inward current was partially (60%) blocked by $100 \mu\text{M}$ bicuculline (BIC). In a multipolar neuron (*b*), the GABA response was completely abolished by $100 \mu\text{M}$ bicuculline ($\text{IC}_{50}=0.4 \mu\text{M}$; not shown). *c*, Cl^- selectivity of bicuculline-resistant GABA responses in bipolar cells. With equal intra- and extracellular Cl^- (145 mM) concentrations (●), the reversal potential of GABA-induced currents was close to 0 mV (mean $-4 \text{ mV} \pm 3 \text{ mV}$; 4 cells). Upon replacement of 130 mM internal Cl^- by equal amounts of gluconate (▲), the reversal potential was shifted to -62 mV (mean $-59 \text{ mV} \pm 4 \text{ mV}$; 5 cells), which is close to the value of -57 mV predicted by the Nernst equation for Cl^- -selective channels. GABA concentration was $20 \mu\text{M}$ in all experiments.

METHODS. Cultures: Retinal slices from 6-day-old Wistar rats were cut at a thickness of $100 \mu\text{m}$ and cultured using the roller-tube technique¹¹ as described¹². Slices were immersed in chicken plasma and thrombin was added to induce clotting. Culture tubes were incubated and rotated in a roller drum at 37°C . The medium contained Eagle's basal medium and Hank's BSS (2:1). Antimitotic drugs (uridine, cytosine arabinoside and fluorodeoxyuridine; $10 \mu\text{M}$ each) were added to inhibit the growth of non-neuronal cells. Rod bipolar cells, which selectively stain with antibodies against protein kinase C¹³, were identified in the inner nuclear layer by fluorescent dye injection¹². Electrophysiological experiments were started after 2–6 weeks *in vitro*. Electrophysiology: The preparation was viewed with Nomarski optics at $640\times$ magnification and continuously perfused (0.5 ml min^{-1}) at room temperature (21 – 25°C). The extracellular bath solution contained (in mM): 137 NaCl , 5.4 KCl , 1.8 CaCl_2 , 1 MgCl_2 , 5 HEPES ($\text{pH } 7.4$). Membrane currents were recorded with whole-cell and outside-out patch methods²⁵, using a HEKA EPC-9 amplifier. Pipettes were coated with silicone elastomer, then filled with a solution containing (in mM): 120 CsCl , 20 TEA-Cl , 1 CaCl_2 , 2 MgCl_2 , 11 EGTA , 10 HEPES ($\text{pH } 7.2$). Pipette resistance was 5 – $7 \text{ M}\Omega$. Lucifer Yellow (potassium salt, 1 mg ml^{-1} ; Sigma) was included in the pipette for visualizing cell morphology. We corrected routinely for the liquid-junction potential⁶; the series resistance of the pipettes (12 – $20 \text{ M}\Omega$) was compensated by up to 80% . Whole-cell current-voltage relations were constructed by ramping the command voltage (100 mV s^{-1}) from the -70 mV holding potential to $+70 \text{ mV}$. Current-voltage relations of GABA-activated responses were obtained by subtracting ramps in the presence and absence of GABA. GABA was applied in the extracellular bath solution to single cells and with a U-tube application system⁷ to excised membrane patches. GABA-activated single-channel currents were studied in outside-out membrane patches (resistance 1 – $30 \text{ G}\Omega$). Data were sampled at 2 kHz and analysed after low-pass filtering at 500 Hz (-3 dB). Half-amplitude threshold analysis was done with a semi-automatic procedure, using the Instrutech



M2-Lab software package. With this system the shortest detectable single-channel event was 0.33 ms ²⁶. Only the most frequently occurring conductance state (main state) was analysed. Single-channel current-voltage (i - V) relations were constructed by determining the average current amplitude at various holding potentials (range -100 mV to 70 mV). Conductance was derived from the slope of the i - V relations by linear regression. Channel openings were defined as bursts of openings²⁶, in which the intraburst intervals did not exceed 2 – 4 ms . Flunitrazepam (gift from Hoffmann-La-Roche) and alphaxalone (Sigma) were prepared in dimethylsulphoxide (DMSO) as 10 mM stock and added to the GABA solution. In control experiments, no effect of 0.01% DMSO on GABA-induced membrane currents was found. Reagents: Bicuculline methiodide (Research Biochem.) was freshly prepared and dissolved directly into extracellular solution; CACA (*cis*-4-aminocrotonic acid) and TACA (*trans*-4-aminocrotonic acid) were purchased from Tocris Neuramin; baclofen and 2-hydroxysaclofen were from Research Biochem.; pentobarbital sodium (Nembutal) was from Serva; picrotoxinin was from Sigma.

pentobarbital and alphaxalone, all potent modulators of GABA_A receptors^{1,2}. In bipolar cells, these drugs did not affect GABA responses recorded in the presence of bicuculline (Fig. 2a; Table 1). Likewise, the Cl^- channel blocker picrotoxinin, (10 – $100 \mu\text{M}$), and the glycine receptor antagonist strychnine ($5 \mu\text{M}$) had little effect (Table 1). In contrast, pronounced effects were observed in multipolar cells containing exclusively bicuculline-sensitive GABA_A receptors (Table 1). Antagonistic action of strychnine has been reported for the GABA_A receptor¹⁴.

The bicuculline-insensitive GABA responses were not mediated by GABA_B receptors. The GABA_B receptor antagonist 2-hydroxysaclofen ($50 \mu\text{M}$), when tested in the presence of bicuculline, did not reduce the GABA response. Consistent with this finding, the GABA_B agonist baclofen ($50 \mu\text{M}$) had no effect either (Fig. 2b). In summary, we detected GABA receptors in the rat retina which are neither of the GABA_A nor GABA_B type.

It has been suggested that folded analogues of GABA may interact selectively with a class of binding sites insensitive to both bicuculline (GABA_A) and baclofen (GABA_B)^{9,10}. We have examined *cis*- and *trans*-4-aminocrotonic acid (CACA and TACA), two GABA analogues of restricted conformation. Inward currents were induced by both compounds, however, with apparently differing activation properties (Fig. 2c, d). The

folded compound CACA ($100 \mu\text{M}$) evoked small but consistent responses, comprising about 10% of the current induced by $20 \mu\text{M}$ GABA. In six cells, where bicuculline blocked the GABA-induced current by $54\% \pm 21\%$, CACA responses were reduced by only $9\% \pm 23\%$, indicating a preference of CACA for bicuculline-insensitive GABA receptors (Fig. 2c). TACA ($20 \mu\text{M}$), the extended GABA analogue, showed no such preference, producing effects similar to those of GABA (Fig. 2d). Hence, the bicuculline- and baclofen-insensitive GABA receptors present in rat retinal bipolar cells may be similar to what have been called¹⁰ GABA_C receptors.

The conductance and gating properties of the novel GABA receptors were distinct from the GABA_A type, although both receptors gate Cl^- selective channels. GABA-activated single-channel currents studied in the presence of $100 \mu\text{M}$ bicuculline revealed small amplitudes of only about 0.5 pA at -70 mV , using symmetrical Cl^- concentrations of 145 mM (Fig. 3a). The conductance obtained from the current-voltage relationship was $7.4 \text{ pS} \pm 0.8 \text{ pS}$ (6 patches). Examination of the gating properties revealed open times of the order of 150 ms (range 120 – 180 ms ; 4 patches). The conductance and open times are clearly different from values obtained for GABA_A receptor channels studied in a variety of cultured central neurons, including rat retinal ganglion cells^{1,6,7,15}. They are also different from GABA_A receptor

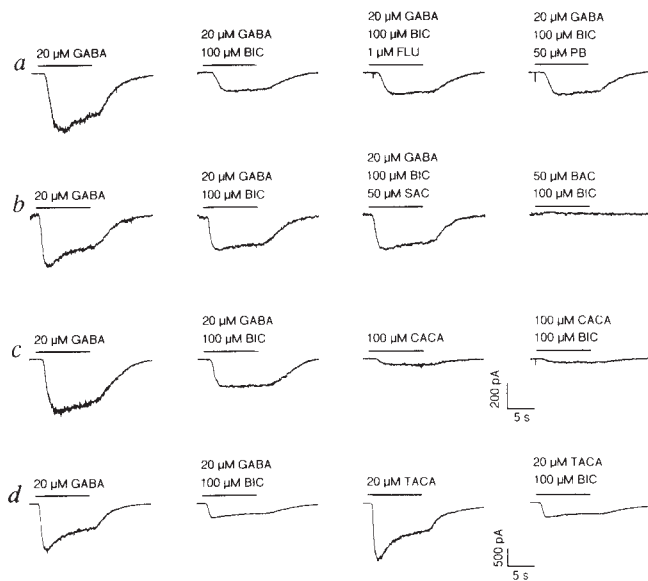


FIG. 2 Pharmacology of bicuculline-insensitive GABA receptors. Whole-cell recordings are shown from four different bipolar cells (a–d) which showed strong GABA (20 μ M) responses in the presence of 100 μ M bicuculline (BIC). a, Bicuculline-insensitive current was only slightly affected by 1 μ M flunitrazepam (FLU) and 50 μ M pentobarbital (PB). b, Bicuculline-resistant current was insensitive to the GABA_B antagonist 2-hydroxysaclofen (SAC) (50 μ M). The GABA_B agonist baclofen (BAC) did not induce any measurable response at 50 μ M. c, CACA (*cis*-4-aminocrotonic acid; 100 μ M) evoked only 7% (range 5–14%; 5 cells) of the peak-current amplitude obtained with 20 μ M GABA. Bicuculline reduced the CACA response by 13%, compared with 47% reduction seen with GABA. d, TACA (*trans*-4-aminocrotonic acid; 20 μ M) produced effects similar to those of GABA (20 μ M). The percentage inhibition by 100 μ M bicuculline was 71% for the GABA response and 73% for the TACA response. Scale bar in c also applies to a and b.

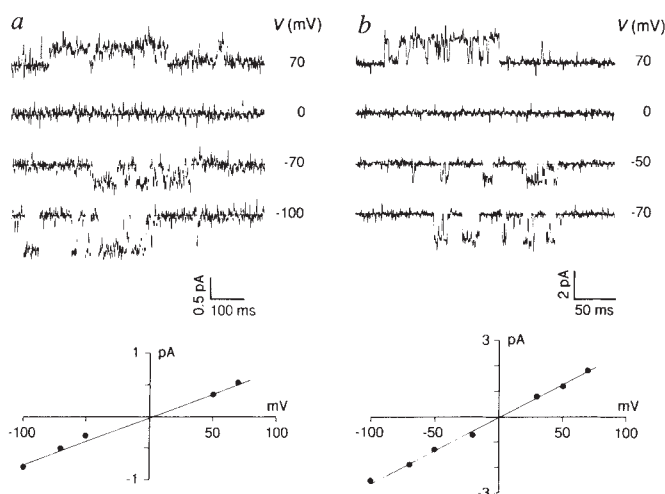


FIG. 3 Single-channel properties of bicuculline-insensitive and sensitive GABA receptors. a, Outside-out patch taken from a bipolar cell. Single-channel activity was started by co-applying 20 μ M GABA and 100 μ M bicuculline to the patch. The holding potentials are indicated on the current traces. Filtered at 200 Hz (–3 dB). The current–voltage relation is a straight line, crossing the abscissa at 0 mV and revealing a single-channel conductance of 7 pS. b, Patch from a bicuculline-sensitive cell, indicating GABA receptors with a single-channel conductance of 27 pS. Filtered at 500 Hz (–3 dB). Note the different scale bars in a and b.

Cl[–] channels in multipolar neurons of the retinal slice culture, for which we obtained a single-channel conductance of 27.1 pS $\pm$ 0.9 pS (16 patches) (Fig. 3b) and a mean open time of 25 ms (range 18–33 ms; 3 patches).

It is very likely that bipolar cells express conventional GABA_A receptors as well as the new type of GABA receptor. Bicuculline-sensitive GABA effects have been reported for rod bipolar cells in rodents only from whole-cell recordings^{16–18}. Consistent with those findings, we could block part of the bipolar GABA response with bicuculline. It should be noted, however, that one study on acutely isolated bipolar cells reports complete blockage of GABA responses by bicuculline¹⁸.

Rod bipolar cells depolarize in response to light and they receive axonal GABA-ergic inputs from amacrine cells via reciprocal and conventional synapses¹³. Whether the novel GABA receptors studied here are involved in these modulatory pathways and what their specific function might be, as opposed to GABA_A receptors in other inhibitory circuits of the central nervous system, remains open.

It is tempting to assume that retinal GABA receptors are assembled from the recently cloned ρ_1 subunit, which is highly expressed in the retina⁴. This peculiar subunit shares only 30–38% homology with other GABA receptor subunits⁴ and forms homo-oligomeric channels when expressed in *Xenopus* oocytes^{4,5}. Although most of the pharmacology of ρ_1 receptors agrees well with our data, there is a marked difference with respect to picrotoxinin sensitivity. The potent action of picrotoxinin at ρ_1 receptor channels (50 per cent inhibitory concentration IC₅₀ < 200 nM) contrasts with the weak antagonistic effects at the novel GABA receptors studied here, and also with the GABA receptors expressed in *Xenopus* oocytes after injecting retinal messenger RNA³. This situation is reminiscent of homo-oligomeric receptors formed upon expressing the glycine receptor α -subunits in *Xenopus* oocytes¹⁹ and in HEK-293 cells^{20,21}, which are strongly blocked by picrotoxinin; but native glycine receptors²² and glycine receptors assembled as α/β hetero-oligomers in HEK cells²¹ show little sensitivity towards the antagonist. Interestingly, the M2 segment that forms the lining of the ρ_1 receptor channel shares more homology with the M2 segment of the glycine receptor α -subunits than with any of the established GABA receptor subunits^{4,23}. Thus it is conceivable that the novel GABA receptors present in the retina are hetero-oligomers assembled from the ρ_1 subunit in combination with another subunit. As glycine receptor β -subunit transcripts are distributed more widely in the central nervous system than those of the glycine receptor α -subunits²⁴, this raises the possibility that the ρ_1 subunit may form picrotoxinin-insensitive channels when coassembled with the glycine receptor β -subunit *in vivo*. □

Received 8 July; accepted 2 November 1992.

- Bormann, J. *Trends Neurosci.* **11**, 112–116 (1988).
- Sivilotti, L. & Nistri, A. *Prog. Neurobiol.* **36**, 35–92 (1991).
- Polenzani, L., Woodward, R. M. & Miledi, R. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4318–4322 (1991).
- Cutting, G. R. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2673–2677 (1991).
- Shimada, S., Cutting, G. & Uhl, G. R. *Molec. Pharmacol.* **41**, 683–687 (1992).
- Bormann, J., Hamill, O. P. & Sakmann, B. *J. Physiol., Lond.* **385**, 243–286 (1987).
- Bormann, J. & Clapham, D. E. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2168–2172 (1985).
- Bormann, J. & Kettenmann, H. *Proc. natn. Acad. Sci. U.S.A.* **85**, 9336–9340 (1988).
- Johnston, G. A. R. et al. *J. Neurochem.* **24**, 157–160 (1975).
- Drew, C. A., Johnston, G. A. R. & Weatherby, R. P. *Neurosci. Lett.* **52**, 317–321 (1984).
- Gähwiler, B. H. *Neuroscience* **11**, 751–760 (1984).
- Feigenspan, A., Bormann, J. & Wässle, H. *Visual Neurosci.* (in the press).
- Wässle, H., Yamashita, M., Greferath, U., Grünert, U. & Müller, F. *Visual Neurosci.* **7**, 99–112 (1991).
- Shirasaki, T., Klee, M. R., Nakaye, T. & Akaike, N. *Brain Res.* **561**, 77–83 (1991).
- Lipton, S. A. *Visual Neurosci.* **3**, 275–279 (1989).
- Karschin, A. & Wässle, H. *J. Neurophysiol.* **63**, 860–876 (1990).
- Suzuki, S., Tachibana, M. & Kaneko, A. *J. Physiol.* **421**, 645–662 (1990).
- Yeh, H. H., Lee, M. B. & Cheun, J. E. *Visual Neurosci.* **4**, 349–357 (1990).
- Schmieden, V., Grenningloh, G., Schofield, P. R. & Betz, H. *EMBO J.* **8**, 695–700 (1989).
- Sontheimer, H. et al. *Neuron* **2**, 1491–1497 (1989).
- Pribilla, I., Takagi, T., Langosch, D., Bormann, J. & Betz, H. *EMBO J.* (in the press).
- Becker, C. M. in *Handbook of Experimental Pharmacology* (eds Herken, H. & Hucho, F.) Vol. 102, 539–575 (Springer, Berlin, 1992).
- Grenningloh, G. et al. *EMBO J.* **9**, 771–776 (1990).
- Malosio, M.-L., Marquière-Pouey, B., Kuhse, J. & Betz, H. *EMBO J.* **10**, 2401–2409 (1991).

25. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch.* **391**, 85–100 (1981).
 26. Colquhoun, D. & Sigworth, F. J. in *Single-Channel Recording* (eds Sakmann, B. & Neher, E.) 191–263 (Plenum, New York, 1983).

ACKNOWLEDGEMENTS. We thank G. Heiss-Herzberger and F. Boij for technical assistance, I. Odenthal for secretarial assistance, and R. Taylor for critically reading the manuscript.

Novel GABA responses from rod-driven retinal horizontal cells

Haohua Qian & John E. Dowling

Department of Cellular and Developmental Biology, Harvard University,
 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

γ -AMINOBUTYRIC acid (GABA) is the main inhibitory neurotransmitter in the central nervous system. Two classes of GABA receptors ($GABA_A$ and $GABA_B$) have been identified. $GABA_A$ receptors are ligand-gated chloride channels that are competitively antagonized by bicuculline, noncompetitively blocked by picrotoxin, and often allosterically modulated by barbiturates and benzodiazepines^{1–3}. $GABA_B$ receptors regulate potassium and calcium channels through G-protein and intracellular second-messenger pathways^{2,3}, are selectively activated by baclofen, and are antagonized by phaclofen and 2-hydroxysaclofen^{4,5}. For some years, evidence has accumulated that there are GABA receptors, especially prominent along visual pathways, which are neither antagonized by bicuculline nor activated by baclofen, but are activated by certain conformationally restricted analogues of GABA, including *cis*-4-aminocrotonic acid (CACA)^{6–9}. These receptors have been designated $GABA_C$ receptors⁹. As yet, membrane current responses from isolated neurons that reflect this novel pharmacology have not been reported, although such responses have been recorded from oocytes injected with retinal messenger RNA^{10–13}. Here we describe a chloride-mediated current response from isolated rod-driven horizontal cells (H4) of the white perch retina that has this novel pharmacology.

Figure 1 shows a typical inward current response recorded from an isolated H4 horizontal cell upon application of 100 μ M GABA to the cell for 30 seconds. The cell was voltage-clamped with a patch pipette and held at -50 mV. Similar responses were observed in all 96 H4 horizontal cells tested, but not in 10 cone-driven horizontal cells. Most cone-driven horizontal cells in white perch are unresponsive to GABA; a small proportion (<25%) show $GABA_A$ receptor-mediated responses (Z. J. Zhou, G. L. Fain and J.E.D. manuscript submitted). To compare the GABA response of the rod-driven horizontal cell with typical $GABA_A$ mediated current responses, a recording from a white perch bipolar cell is also shown in Fig. 1. The responses differ in two respects; first, the response from the bipolar cell showed a dramatic desensitization during the 30-s GABA exposure (in 21 cells tested, a reduction to $37.1 \pm 17.8\%$ of its peak value was observed after 8 s of continuous application of GABA), whereas the response from the horizontal cell showed no desensitization (the response remained at $98.9 \pm 6.7\%$ of its initial value after 8 s of continuous application of GABA; $n = 29$). Second, the H4 horizontal cell response recovered quite slowly after GABA was withdrawn (time constant 15.9 ± 6.0 s; $n = 25$), whereas the response from the bipolar cell quickly returned to the baseline (time constant 1.45 ± 1.12 s, $n = 21$).

Figure 2a shows a typical current-voltage relation for the GABA response recorded from H4 horizontal cells. The positive slope of the line indicates that a membrane conductance increase occurred during the GABA response. The reversal potential for the horizontal cell response was close to 0 mV when chloride was the main extracellular anion (Fig. 2a). Reversal potentials (mean $\pm$ s.d.; $n = 4$) determined with various extracellular chloride concentrations are shown in Fig. 2b. The straight line

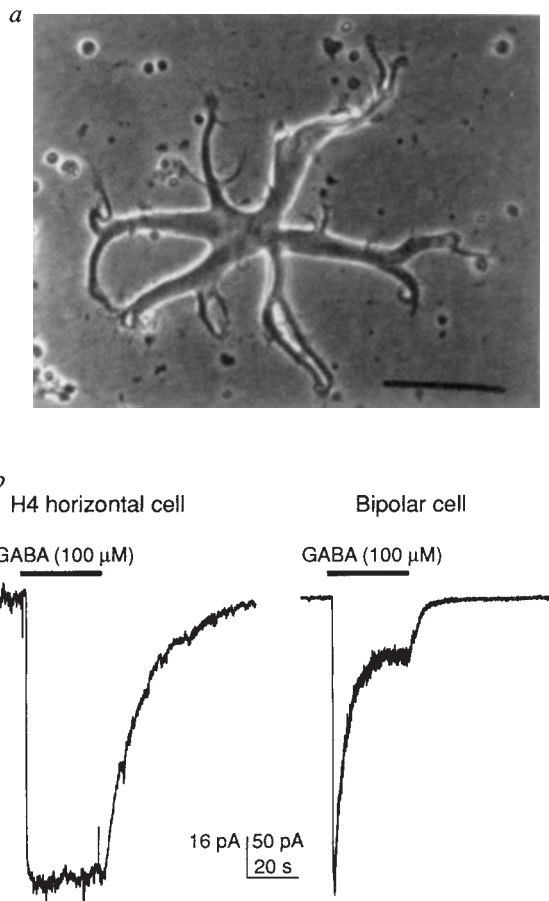


FIG. 1 a, A solitary rod-driven horizontal cell (H4) isolated from the white perch retina. The cell has a flat cell body and several primary processes from which many fine processes extend. Scale bar, 50 μ m. b, GABA responses from an H4 horizontal cell (left) and a bipolar cell (right) isolated from the white perch retina. Both cells were held at -50 mV. GABA was applied onto the cell by bath superfusion, indicated by a bar above each trace. The response from the bipolar cell showed a dramatic desensitization during the application of GABA. The response from the H4 horizontal cell was sustained and did not desensitize during the GABA application. METHODS. Retinal cells were isolated from the white perch (*Roccus americana*) retina by enzymatic and mechanical treatment as described²⁶. Cells were cultured in L-15 medium at room temperature for up to one week. Four subtypes (H1 to H4) of horizontal cells can be identified in culture by their morphology. H1–H3 horizontal cells are cone-driven horizontal cells, whereas H4 horizontal cells are driven exclusively by the rod photoreceptors. The H4 horizontal cell shown in a was in culture for 24 h. Before each experiment the culture medium was replaced by Ringer's solution, which contained (in mM) NaCl (110), KCl (2.5), $CaCl_2$ (2.4), $MgCl_2$ (1.5), glucose (10), HEPES (10), pH 7.7. Whole-cell patch clamp recording techniques were used to monitor membrane currents²⁷. The intracellular solution in the patch pipette contained (in mM) KCl (124), $CaCl_2$ (1), EGTA (11), $MgCl_2$ (2), HEPES (10), Mg-ATP (1). The resistance of a typical patch pipette was ~ 2 M Ω . Series resistance during whole-cell recording was ~ 4 M Ω for horizontal cells and 6 M Ω for bipolar cells. Consequently, a 100 pA current would produce an error of about 0.4 mV in horizontal cell membrane voltage. Recordings were not compensated. Typical input resistance and cell capacitance were about 0.5 G Ω and 250 pF, respectively, for horizontal cells, and 0.5 G Ω and 40 pF for bipolar cells. In some cases, potassium was substituted by caesium to block potassium currents. No difference was found in the GABA-induced currents when the pipettes contained caesium. In most experiments, cells were held at -50 mV and the signals filtered at 30 Hz unless otherwise specified. All values given in the text are expressed as means $\pm$ s.d.

plots the values for the chloride equilibrium potentials calculated from the Nernst equation. The excellent correspondence of the data points to the line indicates the response is mediated by chloride.

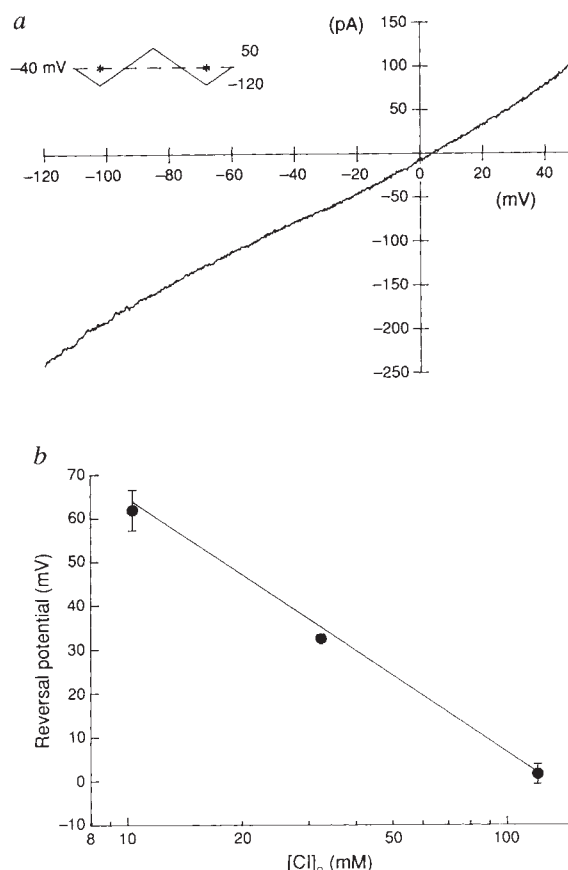


FIG. 2 GABA responses from H4 horizontal cells are mediated by chloride. *a*, Current-voltage relations of the response. Membrane voltage was ramped at 0.5 V s^{-1} according to the protocol shown in the inset. Currents were measured during the time marked by the asterisks; signals were filtered at 1 kHz and averaged 5 times. Current-voltage relations were first recorded in Ringer's and then in Ringer's containing $10 \mu\text{M}$ GABA. The difference between the two averaged records gave the GABA response. *b*, Reversal potentials of GABA responses (means $\pm$ s.d.; $n=4$) determined with various extracellular chloride concentrations. Extracellular chloride was substituted by isothionate. The straight line is the calculated value of the chloride equilibrium potential according to the Nernst equation.

The GABA responses recorded from H4 horizontal cells were bicuculline-insensitive. Figure 3*a, b* and Table 1 show that the GABA responses from H4 horizontal cells were unchanged in the presence of $500 \mu\text{M}$ bicuculline methochloride, whereas the GABA responses from perch bipolar cells were significantly reduced by the same concentration of bicuculline methochloride. (The GABA responsiveness that persists after bicuculline application to bipolar cells may indicate that bipolar cells in white perch possess some bicuculline-resistant GABA receptors.) Picrotoxin ($500 \mu\text{M}$), on the other hand, totally blocked the GABA responses from both the horizontal ($n=12$) and bipolar cells ($n=4$), indicating that both responses are mediated by chloride channels (Fig. 3*c, d*). In addition, neither $10 \mu\text{M}$ diazepam nor $100 \mu\text{M}$ pentobarbital had a significant effect on the GABA responses from horizontal cells (Fig. 3*e, g*; Table 1), whereas both chemicals dramatically enhanced GABA responses from bipolar cells (Fig. 3*f, h*; Table 1). The response was also insensitive to GABA_B receptor agonists and antagonists. Baclofen, a GABA_B agonist, at $100 \mu\text{M}$ did not induce any response from H4 horizontal cells, nor did it alter the GABA current responses ($n=8$). Furthermore, neither $500 \mu\text{M}$ phaclofen ($n=3$) nor $500 \mu\text{M}$ hydroxysaclofen ($n=3$), antagonists of the GABA_B receptors, had any effect on GABA-induced currents in horizontal cells (data not shown). Electrogenic GABA transporters were also excluded as sources of the responses because the GABA-induced current persisted when choline was substituted for extracellular sodium ($n=6$) (Fig. 3*i*).

A conformationally restricted analogue of GABA, *cis*-4-aminocrotonic acid (CACA), elicited a sustained inward current from H4 horizontal cells similar to the GABA response (Fig. 4*a*). The CACA responses, like those of GABA, were blocked by picrotoxin but insensitive to bicuculline ($n=2$). But CACA was less potent than GABA. Dose-response curves for GABA, CACA and muscimol are shown in Fig. 4*b*. The EC_{50} (concentration producing 50% of maximum response) for GABA in

eliciting the response was $1.87 \mu\text{M}$, which is about 10-fold lower than the EC_{50} for GABA at typical GABA_A receptor sites¹⁴⁻¹⁷. Muscimol, a powerful GABA_A agonist, induced only a small sustained current in H4 horizontal cells. The EC_{50} for muscimol ($4.81 \mu\text{M}$) was higher than that for GABA itself, and even at $100 \mu\text{M}$ elicited a response which was only ~42% of that elicited by GABA (Fig. 4*b*). At GABA_A receptors, muscimol is typically more potent than GABA by 2-10 fold¹⁸⁻²⁰.

The GABA responses recorded from H4 horizontal cells of the white perch retina fit well with the properties of bicuculline-resistant GABA responses recorded from the frog optic tectum and guinea-pig superior colliculus^{8,21}; with GABA responses recorded from *Xenopus* oocytes injected with bovine retinal mRNA^{10,11}; and with GABA responses obtained from *Xenopus* oocytes expressing the ρ_1 GABA receptor subunit whose gene has prominent retinal expression^{12,13}. Johnston⁹ has proposed that the receptors mediating such responses be called GABA_C receptors, and Shimada *et al.*¹³ have recently provided arguments supporting this proposal.

TABLE 1 Effect of different drugs on GABA responses in H4 horizontal cells and bipolar cells

Drug	H4 horizontal cells	Bipolar cells
Bicuculline methochloride	$95.8 \pm 7.6, n=10$	$17.8 \pm 2.9^{**}, n=4$
Diazepam	$97.5 \pm 10.8, n=11$	$177.4 \pm 37.2^*, n=4$
Pentobarbital	$108.4 \pm 11.7, n=7$	$586.4 \pm 284.7^*, n=4$

Data are presented as per cent of the GABA response obtained in Ringer's (means $\pm$ s.d.). Drug concentrations were the same as in the experiments shown in Fig. 3.

** $P < 0.001$ compared with 100% control; * $P < 0.05$ compared with 100% control.

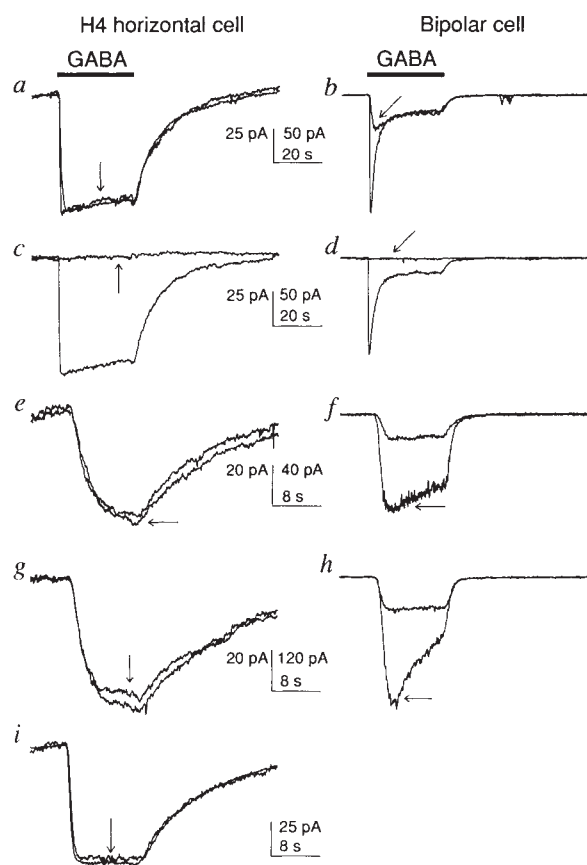


FIG. 3 GABA responses from H4 horizontal cells are pharmacologically distinct from GABA responses of bipolar cells. Arrows point to GABA responses obtained after application of various agents to the cells by bath superfusion. GABA was applied either through a puff pipette filled with 1 mM GABA (*a-d*) or co-applied with the test agent by bath superfusion (*e-i*). *a*, 500 μ M bicuculline methochloride had no effect on the GABA response from the horizontal cell; *b*, 83% of the GABA response from the bipolar cell was blocked by bicuculline methochloride. In five experiments, 10 μ M GABA was co-applied by bath superfusion with 500 μ M bicuculline methochloride. In these experiments as well, no effects on the GABA responses from H4 horizontal cells were noted, whereas in two bipolar cells tested, there was a reduction of $\sim$ 81% in the GABA response. Responses from both the horizontal (*c*) and bipolar cell (*d*) were blocked completely by 500 μ M picrotoxin. GABA responses elicited with 2 μ M GABA were not significantly altered by coapplication of either 10 μ M diazepam (*e*) or 100 μ M pentobarbital (*g*) onto H4 horizontal cells, whereas these drugs at the same concentrations significantly increased the responses of bipolar cells to 10 μ M GABA (*f*, diazepam; *h*, pentobarbital). *i*, 10 μ M GABA responses from an H4 horizontal cell in normal Ringer's and in Ringer's in which choline was substituted for sodium. Traces have been adjusted for a slight baseline shift.

That the GABA responses recorded from H4 horizontal cells of the white perch retina resemble so closely the GABA responses of oocytes expressing only the ρ_1 GABA receptor subunit suggests that the functional GABA receptor of the rod horizontal cells in perch may be homo-oligomeric. By analogy with nicotinic acetylcholine receptors, GABA receptors are thought to be pentamers, consisting of several different subunits. At least 15 different GABA receptor subunits have so far been identified, and the various subunits impart different properties to the receptor²²⁻²⁴. It will be of interest to show unequivocally that GABA receptors made up of only a single subunit are formed in neurons.

The functional significance of these novel GABA receptors is not known. Distal retinal neurons typically respond tonically to prolonged stimuli²⁵, thus GABA receptors that do not desensitize may be of advantage in the outer retina. It may be

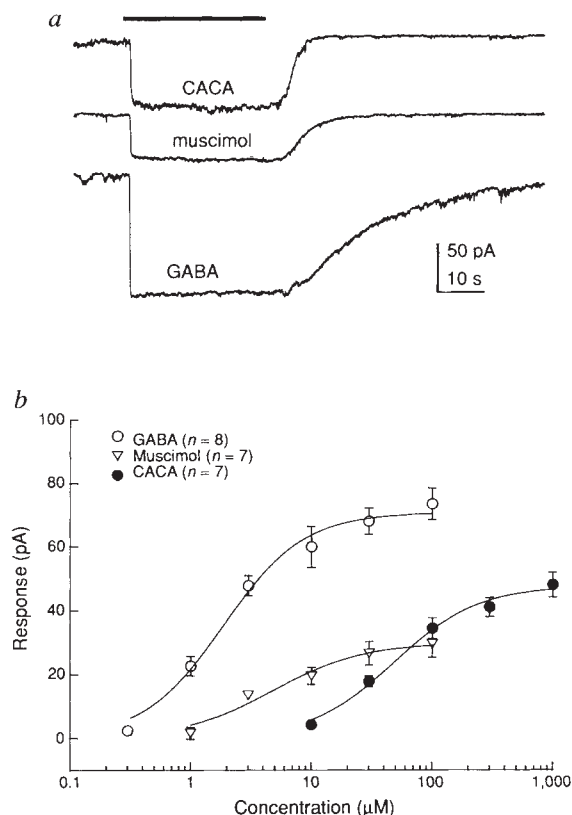


FIG. 4 *a*, Sustained currents elicited in an H4 horizontal cell by bath superfusion of 1 mM *cis*-4-aminocrotonic acid (CACA), 100 μ M muscimol and 100 μ M GABA. *b*, Dose-response curve for GABA, CACA and muscimol. Data are shown as means $\pm$ s.e. The continuous curves are Michaelis-Menten functions, with EC_{50} values of 1.87, 48.5 and 4.81 μ M, and maximum responses of 70.8, 47.5 and 29.8 pA for GABA, CACA and muscimol respectively.

for similar reasons that such receptors are found in higher visual structures and elsewhere in the brain. □

Received 14 September; accepted 2 November 1992.

- Olsen, R. W. *A. Rev. Pharmac. Toxicol.* **22**, 245-277 (1982).
- Bormann, J. *Trends Neurosci.* **11**, 112-116 (1988).
- Sivilotti, L. & Nistri, A. *Prog. Neurobiol.* **36**, 35-92 (1991).
- Hill, D. R. & Bowery, N. G. *Nature* **290**, 149-152 (1981).
- Kerr, D. I. B., Ong, J., Johnston, G. A. R., Abbenante, J. & Prager, R. H. *Neurosci. Lett.* **92**, 92-96 (1988).
- Johnston, G. A. R. et al. *J. Neurochem.* **24**, 157-160 (1975).
- Drew, C. A., Johnston, G. A. R. & Weatherby, R. P. *Neurosci. Lett.* **52**, 317-321 (1984).
- Sivilotti, L. & Nistri, A. *Eur. J. Pharmac.* **164**, 205-212 (1989).
- Johnston, G. A. R. in *Benzodiazepine/GABA Receptors and Chloride Channels. Receptor Biochemistry and Methodology* (eds Olsen, R. W. & Venter, J. C.) Vol. 5, 57-71 (Liss, New York, 1986).
- Polenzani, L., Woodward, R. M. & Miledi, R. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4318-4322 (1991).
- Woodward, R. M., Polenzani, L. & Miledi, R. *Molec. Pharmac.* **42**, 165-173 (1992).
- Cutting, G. R. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2673-2677 (1991).
- Shimada, S., Cutting, G. & Uhl, G. R. *Molec. Pharmac.* **41**, 683-687 (1992).
- Gallagher, J. P., Higashi, H. & Nishi, S. *J. Physiol., Lond.* **275**, 263-282 (1978).
- Akaike, N., Hattori, K., Inomata, N. & Oomura, Y. *J. Physiol., Lond.* **360**, 367-386 (1985).
- Houamed, K. M. et al. *Nature* **310**, 318-321 (1984).
- Itabashi, S., Aibara, K., Sasaki, H. & Akaike, N. *J. Neurophysiol.* **67**, 1367-1374 (1992).
- Mathison, R. D. & Dreifuss, J. J. *Brain Res.* **187**, 476-480 (1980).
- Curtis, D. R., Bornstein, J. C. & Lodge, D. *Brain Res.* **194**, 255-258 (1980).
- Nakagawa, T., Wakamori, M., Shirasaki, T., Nakaye, T. & Akaike, N. *Am. J. Physiol.* **260**, c745-c749 (1991).
- Arakawa, T. & Okada, Y. *Eur. J. Pharmac.* **158**, 217-224 (1988).
- Olsen, R. W. & Tobin, A. J. *FASEB J.* **4**, 1469-1480 (1990).
- Seeburg, P. H. et al. *Cold Spring Harbor Symp. quant. Biol.* **55**, 29-40 (1990).
- Burt, D. R. & Kamatchi, G. L. *FASEB J.* **5**, 2916-2923 (1991).
- Dowling, J. E. *The Retina an Approachable Part of the Brain* (Harvard University Press, Cambridge, Massachusetts, 1987).
- Dowling, J. E., Pak, M. W. & Lasater, E. M. *Brain Res.* **360**, 331-338 (1985).
- Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Archiv.* **391**, 85-100 (1981).

ACKNOWLEDGEMENTS. This research was supported in part by a grant from the National Eye Institute, NIH and Fight for Sight, Research Division of the National Society to Prevent Blindness.

Protein kinase C modulates glutamate receptor inhibition of Ca^{2+} channels and synaptic transmission

Kenton J. Swartz*, Andrew Merritt†, Bruce P. Bean* & David M. Lovinger†

* Department of Neurobiology, Harvard Medical School, Boston, Massachusetts 02115, USA

† Department of Molecular Physiology and Biophysics, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA

FAST synaptic transmission in the central nervous system can be modulated by neurotransmitters and second-messenger pathways. For example, transmission at glutamatergic synapses can be depressed by the metabotropic glutamate receptor^{1,2}, providing autoreceptor-mediated negative feedback. Metabotropic glutamate receptor inhibition of Ca^{2+} channels may contribute to this pathway³⁻⁶. In contrast, stimulation of protein kinase C can enhance excitatory synaptic transmission⁷, whereas both depression and enhancement of Ca^{2+} current have been reported⁸. Here we show that in hippocampal CA3 and cortical pyramidal neurons, activation of protein kinase C enhances current through N-type Ca^{2+} channels and, in addition, dramatically reduces G protein-dependent inhibition of these same channels by the metabotropic glutamate receptor. In parallel experiments on fast excitatory transmission at corticostriatal synapses, kinase C activators were similarly found to reduce the inhibitory effect produced by stimulation of the metabotropic glutamate receptor. The results show that second-to-second control of Ca^{2+} channels by the metabotropic glutamate receptor can itself be modulated on a slower timescale by protein kinase C. These mechanisms may be used in the control of fast excitatory synaptic transmission.

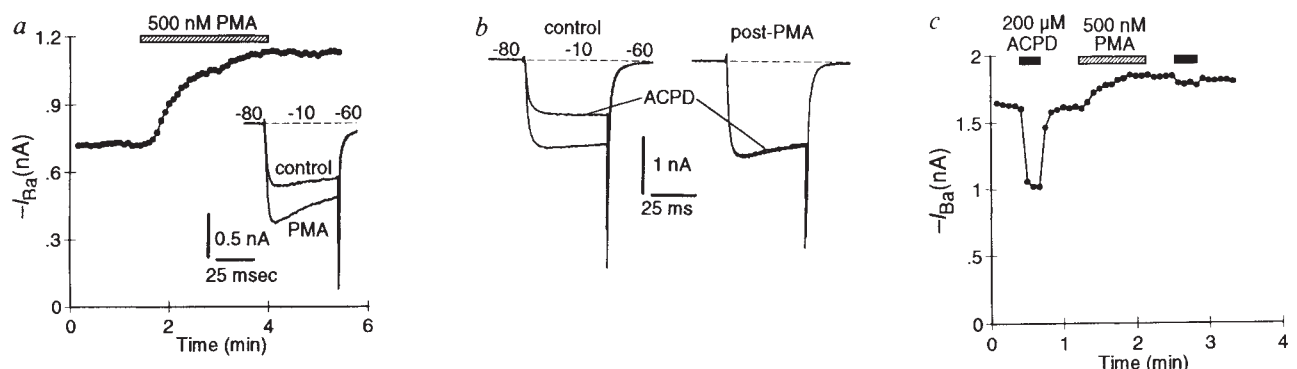


FIG. 1 PMA enhances Ca^{2+} channel current and reduces the suppression of current by a metabotropic glutamate receptor agonist. **a**, Time course of enhancement of peak I_{Ba} 500 nM PMA in a freshly dissociated CA3 pyramidal neuron. Inset shows currents before and 2 min after application of PMA. Numbers above traces indicate holding potential, test potential and tail potential (left to right). In 30 CA3 neurons PMA-increased peak I_{Ba} measured at -10 mV by $28 \pm 2\%$ (18 ± 1 pA pF⁻¹). In another seven CA3 neurons, $1 \mu\text{M}$ PDBu enhanced I_{Ba} (at -10 mV) by $35 \pm 12\%$. **b**, Left, Ca^{2+} channel current from a CA3 pyramidal neuron elicited by depolarization from -80 mV to -10 mV before and 5 s after application of $200 \mu\text{M}$ 1S,3R-ACPD. Right, application of 1S,3R-ACPD after application and washout of PMA (applied for ~ 1 min). **c**, Time course of the effects depicted in **b**. Voltage steps were elicited once every 5 s. 1S,3R-ACPD suppressed I_{Ba} by 34% before PMA application and by 3% afterwards. In six other CA3 pyramidal neurons, $200 \mu\text{M}$ 1S,3R-ACPD inhibited whole-cell Ca^{2+} channel current by $24 \pm 3\%$ (10 ± 2 pA pF⁻¹ absolute current inhibited) before application of 500 nM PMA but by only $3 \pm 1\%$ (1.4 ± 0.5 pA pF⁻¹) following PMA. In seven CA3 neurons $200 \mu\text{M}$ 1S,3R-ACPD suppressed Ca^{2+} channel current by $29 \pm 2\%$ before $1 \mu\text{M}$ PDBu and by $5 \pm 2\%$ following PDBu.

Figure 1 illustrates the effects of phorbol-12-myristate-13-acetate (PMA) on voltage-dependent Ca^{2+} channel current in pyramidal neurons acutely isolated from the CA3 region of the rat hippocampus. Application of PMA (500 nM) increased whole-cell Ba^{2+} current within seconds and took ~ 1 –2 min to reach steady-state (Fig. 1a, c).

Ca^{2+} channel current in hippocampal and cortical neurons can be inhibited by stimulation of metabotropic glutamate receptors³⁻⁶. 1-Aminocyclopentane-1S,3R-dicarboxylic acid (1S,3R-ACPD) is a selective agonist at metabotropic glutamate receptors^{9,10} that rapidly and reversibly suppresses N-type Ca^{2+} channels in CA3 pyramidal neurons⁴. In addition to enhancing Ca^{2+} channel current, PMA dramatically reduced the ability of metabotropic glutamate receptor agonists to modulate Ca^{2+} channels (Fig. 1b, c). In six CA3 pyramidal neurons, a saturating concentration of 1S,3R-ACPD ($200 \mu\text{M}$) reduced the Ba^{2+} current (I_{Ba}) by $24 \pm 3\%$ before application of PMA, but by only $3 \pm 1\%$ after treatment with PMA for ~ 1 min. This result is consistent with phorbol ester effects previously noted with quisqualate applied to GTP- γ S-loaded neurons³.

Further experiments suggested that the actions of PMA are due to activation of kinase C. The inactive analogue 4- α -PMA had no effect on either I_{Ba} or inhibition by 1S,3R-ACPD (Fig. 2a, c). Including a pseudosubstrate inhibitor of protein kinase C (PKC 19-36)¹¹ in the internal solution prevented the effects of PMA on both Ba^{2+} current amplitude and modulation by 1S,3R-ACPD (Fig. 2b, c). The actions of PMA were mimicked by another kinase C activator (phorbol-12,13-dibutyrate (PDBu)) (Fig. 1 legend).

The enhancement of Ca^{2+} channel current by kinase C is mediated, at least in part, by effects on N-type Ca^{2+} channels. A saturating concentration ($10 \mu\text{M}$) of the N-type channel blocker ω -conotoxin fraction GVIA (ω -CgTx)¹²⁻¹⁵ inhibited more current in cells exposed to PMA than in control cells (Fig. 3). The increase in ω -CgTx-sensitive current (from 18 ± 1 pA pF⁻¹ in control cells to 28 ± 3 pA pF⁻¹ after PMA) accounted for $\sim 60\%$ of the total increase (17 ± 2 pA pF⁻¹) in current produced by PMA. In CA3 cells pre-exposed to $10 \mu\text{M}$ ω -CgTx,

METHODS. Pyramidal neurons were freshly dissociated from the CA3 region of hippocampal slices ($400 \mu\text{m}$) obtained from 7–21-day-old Long-Evans rats as previously described⁴. Whole-cell voltage clamp recordings³⁵ were obtained using pipettes with resistances ranging from 0.5 to 5 M Ω when filled with the internal solution consisting of (in mM): 117 TEA-Cl, 4.5 MgCl₂, 9 HEPES, 9 EGTA, 14 creatine phosphate-Tris, 4 Mg-ATP, 0.3 GTP-Tris, pH 7.4. The external recording solution contained (in mM): 25 BaCl₂, 145 TEA-Cl, 10 HEPES, pH 7.4. Phorbol esters were diluted into the external solution from 2 mM stock solutions in dimethyl sulphoxide (DMSO); final concentrations of DMSO were $<0.05\%$. Drug solutions were applied by gravity-driven perfusion from a linear array of 12 microcapillary tubes (exchange rate <250 ms). Whole-cell currents recorded with a List EPC-7 patch clamp amplifier were filtered at 2 kHz (8-pole Bessel low-pass), digitized and stored using a BASIC-FASTLAB analog-digital interface and software (Indec Systems, Sunnyvale, CA). Leak and capacitive currents were subtracted using linearly scaled currents elicited by a step from -80 to -90 mV. All experiments were done at room temperature (22 – 25°C). All statistical results are given as mean $\pm$ s.e.m.

the enhancement produced by PMA was much smaller than control cells (Fig. 3). These results show that about 60–70% of the current enhanced by PMA flows through N-type Ca^{2+} channels sensitive to ω -CgTx. The remainder of the PMA-enhanced current might arise from L-type Ca^{2+} channels (whose activity in cell-attached patches can be increased by phorbol esters^{16–18}), P-type channels, or N-type channels incompletely blocked by ω -CgTx.

We obtained similar results using pyramidal neurons acutely isolated from lower layers of cerebral cortex overlying the striatum, some of which give rise to the presynaptic terminals of corticostriatal synapses. PMA (500 nM) enhanced current

elicited by a step to -10 mV by $51 \pm 13\%$ (27 ± 4 pA pF⁻¹; $n = 4$). In cells first exposed to $10 \mu\text{M}$ ω -CgTx, there was enhancement by PMA of only $9 \pm 3\%$ (5 ± 0.4 pA pF⁻¹; $n = 3$), consistent with PMA enhancing primarily the N-type current. As in hippocampal neurons, $200 \mu\text{M}$ 1S,3R-ACPD inhibited Ca^{2+} channel current, and the current inhibited by 1S,3R-ACPD was largely, but not entirely, N-type current⁴; 1S,3R-ACPD inhibited 13 ± 2 pA pF⁻¹ ($20 \pm 3\%$) of current before and 5 ± 2 pA pF⁻¹ ($7 \pm 2\%$) of current after exposure to $10 \mu\text{M}$ ω -CgTx GVIA; in these 5 cells, ω -CgTx blocked $34 \pm 7\%$ of the total current. Exposure to 500 nM PMA greatly suppressed the inhibition by $200 \mu\text{M}$ 1S,3R-ACPD, from $20 \pm 3\%$ in control to $3.4 \pm 0.3\%$ after treatment with 500 nM PMA ($n = 4$).

A major role of voltage-dependent Ca^{2+} channels in the brain is to couple the presynaptic action potential to transmitter release. N-type Ca^{2+} channels mediate such coupling at some (but not all) synapses^{19–21}. We tested whether the actions of

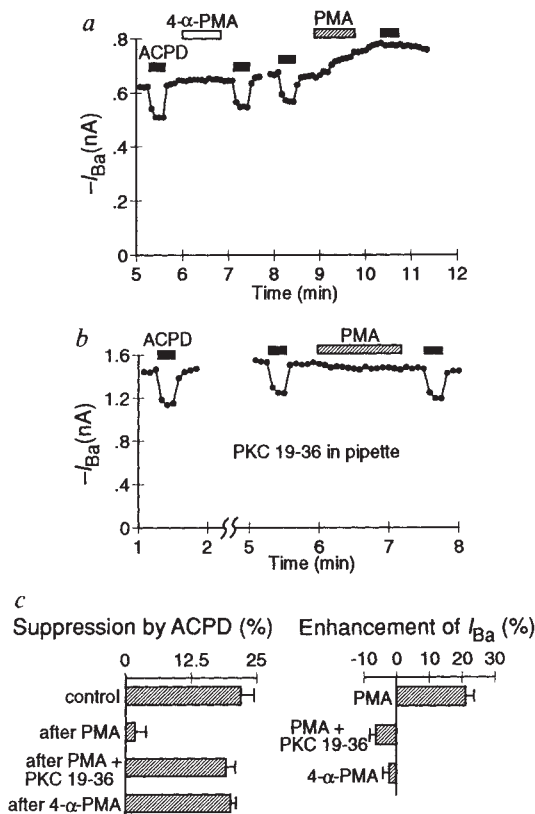


FIG. 2 The effects of PMA are mediated by protein kinase C. *a*, Enhancement of Ca^{2+} channel current and reduction of 1S,3R-ACPD ($200 \mu\text{M}$) suppression by 500 nM 4- β -PMA (PMA) but not 500 nM 4- α -PMA. Voltage steps to -10 mV (50 ms long) were elicited every 5 s from a holding potential of -80 mV. *b*, PKC 19-36 ($200 \mu\text{M}$ in pipette) blocked enhancement of I_{Ba} and reduction of 1S,3R-ACPD suppression by PMA. 1S,3R-ACPD was tested 1 min after obtaining the whole-cell configuration; PKC 19-36 was allowed to diffuse into the cell for 5 min before testing PMA for effects on I_{Ba} and its suppression by 1S,3R-ACPD. Pipette resistance was 0.8 M Ω with an initial series resistance of 4 M Ω (of which 70% was electronically compensated to reduce voltage errors) that decreased slightly during the experiment. Voltage pulse protocol as in *a*. Time axes in both *a* and *b* are time since obtaining whole-cell configuration. *c*, Pooled results showing the effects of 4- α -PMA (500 nM), PMA (500 nM) and PKC 19-36 on Ca^{2+} channel currents and their modulation by 1S,3R-ACPD. PMA was tested in 5 cells with $20 \mu\text{M}$ PKC 19-36 in the pipette and exchange allowed to occur for ~ 20 min. Another five cells were tested where PKC 19-36 was included in the pipette at a concentration of $200 \mu\text{M}$. In these neurons, after only 5 min of exchange, the effects of PMA on I_{Ba} and its suppression by 1S,3R-ACPD were tested. Results with $20 \mu\text{M}$ (20 min exchange) and $200 \mu\text{M}$ (5 min exchange) were similar and were combined in the bar graph shown. Twelve control cells were examined with exchange between pipette and cytoplasm allowed for 5 to 20 min (eight cells 5 – 10 min and 4 cells >20 min). Control and PKC 19-36 cells were tested alternately. 4- α -PMA along with PMA was tested in six cells where control ACPD suppression was larger than 15% . All experiments were done using patch pipettes with resistances of 0.5 – 2 M Ω to allow for rapid exchange between pipette solutions and cytoplasm.

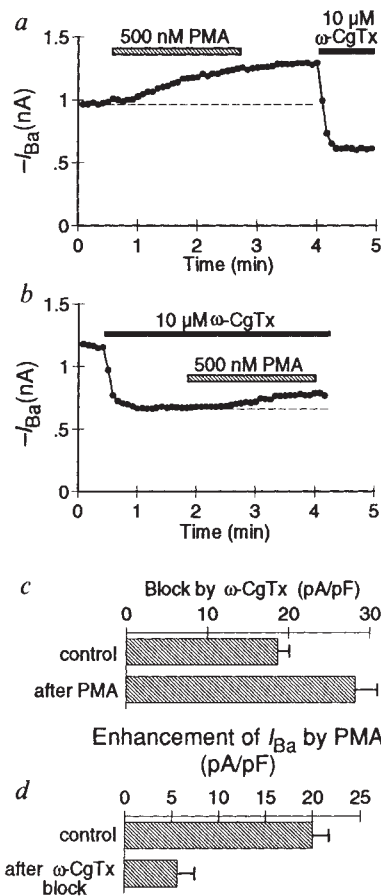


FIG. 3 Phorbol esters enhance N-type Ca^{2+} channel current. *a*, Block by $10 \mu\text{M}$ ω -CgTx (GVIA) after enhancement of I_{Ba} by 500 nM PMA. Voltage steps from -80 mV to -10 mV were elicited every 5 sec. *b*, Enhancement of I_{Ba} after block by ω -CgTx. Same voltage protocol and stimulation frequency as in *a*. *c*, Pooled data showing amount of current (normalized to cell capacitance) blocked by ω -CgTx in control cells ($n = 11$) and in cells following application of 500 nM PMA for ~ 2 min ($n = 11$). Cells were studied only if I_{Ba} was >500 pA. Control current and membrane capacitance were similar in control cells (830 ± 55 pA and 15 ± 1 pF) and PMA-treated cells (828 ± 67 pA and 13 ± 1 pF). *d*, Pooled data showing amount of current enhanced by 500 nM PMA in control cells ($n = 9$) and in cells following block by $10 \mu\text{M}$ ω -CgTx ($n = 6$). Enhancement by PMA was tested alternately in control and ω -CgTx treated cells. Control currents and membrane capacitance were similar in these two groups of cells (850 ± 114 pA and 13 ± 1 pF for control cells and 890 ± 159 pA and 14 ± 3 pF for ω -CgTx treated cells). When expressed as per cent of initial control current, PMA enhanced current by $34 \pm 4\%$ under control conditions and by $11 \pm 4\%$ following ω -CgTx block.

kinase C activation on Ca^{2+} channels might be paralleled by effects on excitatory synaptic transmission at corticostriatal synapses. Single pulse stimuli given within a striatal slice produce monosynaptic excitatory responses which can be recorded as a population spike during field potential recording²² or as an excitatory post-synaptic potential (e.p.s.p.) or excitatory post-synaptic current (e.p.s.c.) during whole-cell recording¹. Transmission at corticostriatal synapses is reversibly suppressed by stimulation of metabotropic glutamate receptors with ACPD, probably through a presynaptic mechanism¹. Racemic ACPD ($\pm$ -trans-ACPD; 50 μM) reduced the population spike amplitude by >30% ($n=70$) and the e.p.s.p. amplitude by >50% ($n=50$) in all slices examined. Transmission is at least partially mediated by N-type Ca^{2+} channels; in whole-cell recordings, 1 μM ω -CgTx reduced e.p.s.p. amplitudes by $69 \pm 4\%$ ($n=4$).

Phorbol-12,13-diacetate (PDAC) and PDBu enhanced excitatory transmission in striatal slices (Fig. 4), consistent with previous findings at hippocampal synapses⁷. In addition, both PDAC and PDBu reduced the ability of $\pm$ -trans-ACPD to suppress excitatory synaptic transmission (Fig. 4c–f). In many cases, where lower concentrations of phorbol esters failed to produce significant enhancement of population spike or e.p.s.p. amplitudes, they still dramatically reduced ACPD suppression of synaptic transmission (Fig. 4c). Consistent with mediation by kinase C, the inactive analogue 4- α -PMA had little effect and the protein kinase inhibitor H-7 attenuated the effects of 1 μM PDBu (Fig. 4 legend).

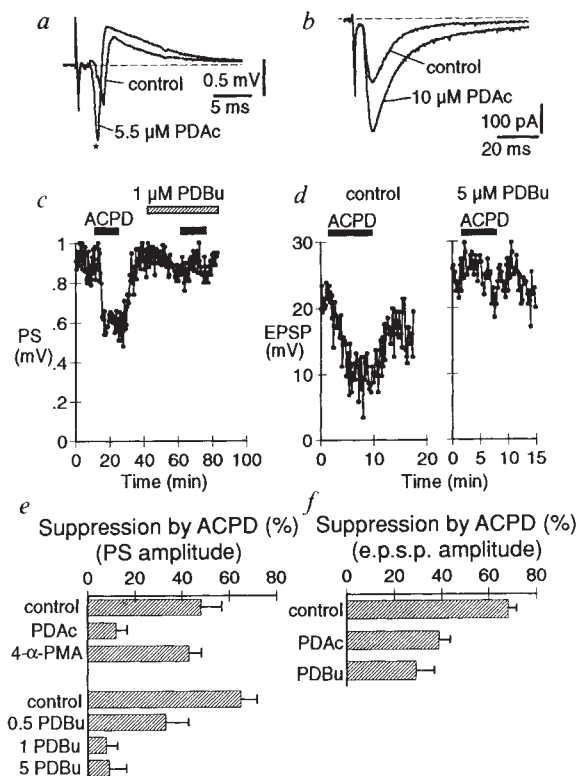
These experiments demonstrate parallels between the modulation of Ca^{2+} channels and of excitatory synaptic transmission:

enhancement by kinase C, inhibition by the metabotropic glutamate receptor, and suppression of that inhibition by kinase C. In particular, N-type Ca^{2+} channels are targeted both by glutamate receptors and by kinase C and contribute to transmission at corticostriatal synapses. Our results do not rule out the involvement of K^{+} channels²³ or vesicular release machinery²⁴ in the actions of the metabotropic glutamate receptor or protein kinase C on synaptic transmission. The disruption of metabotropic glutamate receptor suppression of synaptic transmission is reminiscent of phorbol ester disruption of presynaptic inhibition by GABA_B and adenosine receptors in hippocampus^{25,26}. Preliminary results show that kinase C activation dramatically reduces inhibition of Ca^{2+} channels by GABA_B and adenosine receptors in CA3 and cortical pyramidal neurons (K.J.S., unpublished observations). More generally, activation of kinase C apparently inhibits a variety of intracellular signalling pathways, including those involving phospholipase C^{27–30}, adenylate cyclase³¹ and tyrosine kinase³². Activation of kinase C also disrupts substance P inhibition of M current³³.

Protein kinase C has been proposed to mediate transmitter inhibition of Ca^{2+} channels in some instances^{8,34}, but our results show that in hippocampal neurons, where ACPD inhibition of Ca^{2+} channels is fast, G protein-mediated and membrane delimited⁴, stimulation of kinase C actually enhances Ca^{2+} channel current and reduces ACPD inhibition. Furthermore, metabotropic glutamate receptor inhibition of Ca^{2+} channel current is not prevented by protein kinase C inhibitors. It remains to be determined whether the effects of kinase C are exerted on transmitter receptors, G proteins or Ca^{2+} channels, and how

FIG. 4 Phorbol esters enhance excitatory synaptic transmission and reduce suppression of transmission by a metabotropic glutamate receptor agonist. **a**, Field potentials recorded before and 20 min after superfusion of a striatal slice with 5.5 μM PDAC. The asterisk indicates the population spike (PS). PDAC (5.5 μM) increased PS amplitude by $26 \pm 7\%$ ($n=9$), 1 μM PDAC increased PS amplitude by $14 \pm 6\%$ ($n=6$) and 5 μM PDBu increased PS amplitude by $18 \pm 3\%$ ($n=4$). In five slices pretreated with the protein kinase inhibitor H-7 (200 μM H-7 for 30 min), 1 μM PDBu had no effect on population spike amplitude ($98 \pm 7\%$ of baseline). **b**, Whole-cell excitatory post-synaptic currents recorded before and 8 min after application of 10 μM PDAC. Holding potential was -70 mV. PDAC increased e.p.s.c. amplitude during whole-cell recording by $41 \pm 16\%$ in four slices. **c**, Time course for suppression of PS amplitude by $\pm$ -trans-ACPD (50 μM) before and after application of 5 μM PDBu. Stimulation frequency was 0.05 Hz. **d**, Time course for $\pm$ -trans-ACPD inhibition of e.p.s.p. in one cell under control conditions and in another cell in the same slice 35 min after application of 1 μM PDBu. Stimulation frequency was 0.1 Hz. **e**, Pooled data for the effects of phorbol esters on $\pm$ -trans-ACPD (50 μM) suppression of PS amplitude. In the first series of experiments $\pm$ -trans-ACPD was tested before and after application of 5.5 μM PDAC in five slices. In four slices $\pm$ -trans-ACPD was tested after incubation with 5.5 μM 4- α -PMA. In a second series of experiments $\pm$ -trans-ACPD was tested before and after incubation with either 0.5 μM ($n=4$), 1 μM ($n=6$) or 5 μM ($n=4$) PDBu. $\pm$ -trans-ACPD (50 μM) reduced PS amplitude by $38 \pm 5\%$ ($n=5$) after incubation of slices with 1 μM PDBu plus 200 μM H-7. **f**, Pooled data for the effects of phorbol esters on $\pm$ -trans-ACPD (50 μM) suppression of e.p.s.p. amplitudes obtained with whole-cell recording. $\pm$ -trans-ACPD was tested under control conditions ($n=12$) or after incubation with 5.5 μM PDAC ($n=7$) or 5 μM PDBu ($n=5$).

METHODS. Striatal slices (400 μm ; coronal sections) were prepared from 3–4-week-old rats for whole-cell experiments or >2-month-old rats for extracellular recordings as previously described⁴ and bathed in a solution containing (in mM): 124 NaCl, 4.5 KCl, 2 CaCl_2 , 1.5 MgCl_2 , 26 NaHCO_3 , 1.2 NaH_2PO_4 , and 10 glucose, pH 7.4 with 95% O_2 , 5% CO_2 , 295–300 mOsm. Slices were superfused with this solution at a rate of 1.8 ml min^{-1} . A stimulating electrode, placed in the centre of the striatum, was used with a square wave protocol given by an S88 stimulator and PSU6 optical isolation unit (Grass Inst. Inc.). For extracellular recordings, micropipettes (<1 M Ω) filled with 0.9% saline, were placed 1 mm ventral to the stimulation site within the striatal slice. Pipette position was optimized by recording responses to single stimuli (0.02 to 0.2 ms at 0.5 mA) and setting the depth to a position where the largest synaptically evoked responses were observed. Potentials were amplified (1,000-fold) by a differential AC amplifier, low-pass filtered at 5 kHz and digitized at <125 kHz using pClamp analog-



digital interface and software (Axon Instruments, Foster City, CA). For whole-cell patch clamp recordings, pipettes (2–3 M Ω) contained (in mM): 125 KMeSO₄, 15 KCl, 1 MgCl_2 , 1.1 EGTA, 10 HEPES and 2 Mg-ATP (pH 7.4 with KOH and 280 mOsm with sucrose). Synaptic responses were recorded either in current-clamp at the resting membrane potential, or in voltage-clamp at the desired membrane potential using an Axopatch 1-D patch-clamp amplifier. Responses were filtered and digitized as above for extracellular recordings. Drugs were applied by bath application. All experiments were done at room temperature (20–21 $^{\circ}\text{C}$).

this control pathway is used in specific instances of synaptic plasticity. □

Received 24 July; accepted 10 November 1992.

1. Lvinger, D. M. *Neurosci. Lett.* **129**, 17–21 (1991).
2. Baskys, A. & Malenka, R. C. *J. Physiol., Lond.* **444**, 687–701 (1991).
3. Lester, R. A. & Jahr, C. E. *Neuron* **4**, 741–749 (1990).
4. Swartz, K. J. & Bean, B. P. *J. Neurosci.* **12**, 4358–4371 (1992).
5. Sayer, R. J., Schwandt, P. C. & Crill, W. E. *J. Neurophysiol.* **68**, 833–842 (1992).
6. Sahara, Y. & Westbrook, G. L. *Soc. Neurosci.* **17**, 1168 (1991).
7. Malenka, R. C., Madison, D. V. & Nicoll, R. A. *Nature* **321**, 175–177 (1986).
8. Anwyl, R. *Brain Res. Rev.* **16**, 265–281 (1991).
9. Palmer, E. et al. *Eur. J. Pharmac.* **166**, 585–587 (1988).
10. Irving, A. J. et al. *Eur. J. Pharmac.* **186**, 363–365 (1990).
11. House, C. & Kemp, B. E. *Science* **238**, 1726–1728 (1987).
12. Olivera, B. M. et al. *Science* **230**, 1338–1343 (1985).
13. Plummer, M. R., Logothetis, D. E. & Hess, P. *Neuron* **2**, 1453–1463 (1989).
14. Aosaki, T. & Kasai, H. *Pflügers Arch.* **414**, 150–156 (1989).
15. Regan, L. J., Sah, D. W. Y. & Bean, B. P. *Neuron* **6**, 269–280 (1991).
16. Lipscombe, D., Bley, K. & Tsien, R. W. *Soc. Neurosci.* **14**, 153 (1988).
17. Madison, D. V. *Soc. Neurosci.* **15**, 16 (1989).
18. O'Dell, T. J. & Alger, B. E. *J. Physiol., Lond.* **436**, 739–767 (1991).
19. Hirning, L. D. et al. *Science* **239**, 57–61 (1988).
20. Kamiya, H., Sawada, S. & Yamamoto, C. *Neurosci. Lett.* **91**, 84–88 (1988).
21. Horne, A. L. & Kemp, J. A. *Br. J. Pharmac.* **103**, 1733–1739.
22. Cordingley, G. E. & Weight, F. F. *Br. J. Pharmac.* **88**, 847–856 (1986).
23. North, R. A. *Br. J. Pharmac.* **98**, 13–28 (1989).
24. Scholz, K. P. & Miller, R. J. *Neuron* **8**, 1139–1150 (1992).
25. Dutar, P. & Nicoll, R. A. *Neuron* **1**, 585–591 (1988).
26. Thompson, S. M., Haas, H. L. & Gähwiler, B. H. *J. Physiol., Lond.* **451**, 347–363 (1992).
27. Leeb-Lundberg, L. M. F. et al. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5651–5655 (1985).
28. Pollock, W. K. & MacIntyre, D. E. *Biochem. J.* **234**, 67–73 (1986).
29. Schoepp, D. D. & Johnson, B. G. *Biochem. Pharmac.* **37**, 4299–4305 (1988).
30. Aramori, I. & Nakanishi, S. *Neuron* **8**, 757–765 (1992).
31. Katada, T. et al. *Eur. J. Biochem.* **151**, 431–437 (1985).
32. Bollag, G. E. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5822–5824 (1986).
33. Bosma, M. M. & Hille, B. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2943–2947 (1989).
34. Rane, S. G. et al. *Neuron* **3**, 239–245 (1989).
35. Hamill, O. P. et al. *Pflügers Arch.* **391**, 85–100 (1981).

ACKNOWLEDGEMENTS. This work was supported by the NIH NIAAA and the Mahoney Institute.

Chemomechanical cycle of kinesin differs from that of myosin

Laura Romberg* & Ronald D. Vale†‡

Department of *Biochemistry and †Pharmacology, University of California, San Francisco, California 94143, USA

MOTOR proteins move unidirectionally along cytoskeletal polymers by coupling translocation to cycles of ATP hydrolysis. The energy from ATP is required both to generate force and to dissociate the motor-filament complex in order to begin a new chemomechanical cycle^{1,2}. For myosin, force production is associated with phosphate release following ATP hydrolysis, whereas dissociation of actomyosin is tightly coupled to the binding of ATP³. Dynein, a microtubule motor, uses a similar cycle⁴, suggesting that all cytoskeletal motors might operate by a common mechanism. Here we investigate kinesin's chemomechanical cycle by assaying microtubule movement by single kinesin molecules when intermediate states in the hydrolysis cycle are prolonged with ATP analogues or inhibitors. In contrast to myosin and dynein, kinesin with bound ADP dissociates from microtubules during translocation, whereas kinesin with unhydrolysed nucleotide remains tightly associated with the polymer. These findings imply that kinesin converts ATP energy into mechanical work by a pathway distinct from that of myosin or dynein.

Single kinesin molecules attached to a microscope slide can translocate microtubules for several microns at velocities identical to those produced by multiple motors^{5,6}. As diffusion would separate a detached microtubule from the two-headed kinesin molecule within a millisecond, kinesin must remain bound to the microtubule for most of the ATPase cycle (~50 ms in total⁷).

But each head must dissociate briefly when moving to a new tubulin binding site. To determine when this occurs in the hydrolysis cycle, ATP analogues and inhibitors were used to prolong different chemical intermediates (no nucleotide, ATP, ADP plus hydrolysed P_i (ADP·P_i) and ADP). Agents that prolong kinesin intermediates that are weakly bound should cause microtubules to dissociate more frequently from single kinesin molecules. Those that extend a strongly bound intermediate should allow single motors to move microtubules for many microns, albeit at slower rates. In addition, prolonging a strong binding state should increase the drag kinesin imposes on microtubules, thus causing transport by multiple motors to be slower than that produced by a single motor⁸.

To assay single motor motility, kinesin was adsorbed onto a surface at a low density⁵. Microtubules that pivoted about single attachment points and travelled distances shorter than their

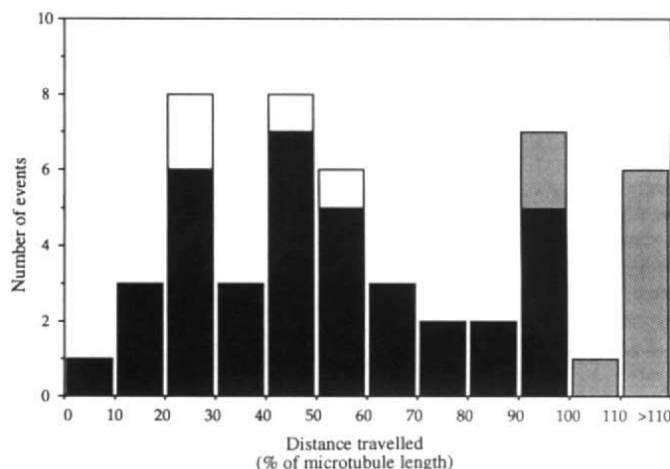


FIG. 1 Microtubule motility on surfaces sparsely coated with kinesin; criteria for establishing movement by single motors. This histogram shows how far microtubules travel before dissociating (expressed as a percentage of their length). Translocation events are subdivided according to whether microtubules moved in a straight line (open bars) or pivoted about one (solid bars) or two successive (shaded bars) points of attachment to the surface. Microtubules that moved further than their length always had a second pivot point, suggesting an interaction with a second kinesin motor. Only those microtubules with a single point of attachment and which did not move beyond their length were considered to have interacted with a single kinesin. Video observations indicate that the ends of microtubules in solution frequently contact the surface; thus initial attachment to motors may not occur at random along the length of a microtubule.

METHODS. Kinesin was purified from squid lobes by microtubule affinity²⁰ followed by sucrose gradient sedimentation (5–20% sucrose prepared in 80 mM PIPES, 1 mM MgCl₂, 1 mM EGTA; spun at 200,000g for 12 h). Motility assays were performed as in ref. 5 with the following modifications. The surface of a microscope perfusion chamber was precoated for 2 min with casein (5 mg ml⁻¹ in 10 mM Tris, pH 7.7, 150 mM NaCl) instead of cytochrome c/tubulin. A solution of 80 mM PIPES, pH 6.8, 1 mM EGTA, 5 mM MgCl₂ was used for all subsequent dilutions and for assaying motility. Kinesin was introduced into the microscope chamber at a concentration of 4 µg ml⁻¹ for the multiple motor assay, and 50 ng ml⁻¹ for the single motor assay, yielding surface densities of 560 and 7 molecules per µm² respectively⁵. By applying Poisson statistics to the probability that a microtubule finds a second motor during translocation (moving 4 µm on average), we calculated that there was an average of one motor per 18 µm along the microtubule path in the single motor assay. Therefore in multiple motor assays about 20 motors were estimated to be attached to each 5 µm long microtubule. Before introduction into the microscope chamber, microtubules stabilized with 10 µM taxol (15 µg ml⁻¹ in the multiple motor assay and 90 µg ml⁻¹ in the single motor assay) were sheared with a 27-gauge needle (average length of 4–7 µm in different experiments). Microtubule motility was observed using dark-field optics. Velocities and distances translocated by single motors were determined by measuring the distance from the ends of microtubules (at the start and end of a translocation event) to the node about which they pivoted. All translocation events occurring within a 1,200 µm² field of view were measured.

‡ To whom correspondence should be addressed.

TABLE 1 Microtubule interaction times and velocities with different nucleotides

Nucleotide	Number of releases*	Dissociation from single motors			Velocity ($\mu\text{m s}^{-1}$)†		
		Seconds (μm) bound to MTs	Seconds bound per release	μm travelled per release	One motor	Multiple motors	Ratio
1 mM ATP	17	186 (152)	11	9	0.82 ± 0.03	0.75 ± 0.01	1.1
2 μM ATP	29	4,297 (93)	148	3	0.022 ± 0.001	0.011 ± 0.001	2
250 μM ATP- γS ‡	25	4,149 (67)	166	3	0.016 ± 0.001	0.012 ± 0.0003	1.4
1 mM ATP/150 μM ADP	39	212 (153)	5	4	0.72 ± 0.02	0.67 ± 0.01	1.1
1 mM ATP/250 μM ADP	—	0 (0)§	—	—	—	0.55 ± 0.01	—

Nucleotides were purchased from Boehringer Mannheim. ATP- γS was further purified by chromatography on a Mono-Q column (Pharmacia) with a linear 0–0.6 M triethylamine gradient. The final product was analysed by 0.3 M ammonium bicarbonate isocratic elution. It contained 0.5% ATP and 1% ADP; the latter increased to 5% over a period of one month. The ATP- γS , used shortly after preparation, was assayed with AP5A (P^1P^5 -di(adenosine-5')-pentaphosphate; 100 μM) to inhibit any dinucleotide kinase that may have been present. An ATP-regenerating system (0.1 mg ml^{-1} creatine kinase and 1 mM phosphocreatine) was included in the assay with 2 μM ATP.

* A translocating microtubule is considered to dissociate from a single motor if it is released into solution before its trailing end reaches the pivot point where the kinesin molecule is presumably located. For each condition, data were taken from multiple microtubules on at least three slides.

† The translocation velocities produced by one and multiple motors are expressed separately and as a ratio. Microtubules in multiple motor assays are estimated to be bound by ~20 kinesins (see Fig. 1). The means and standard errors from at least 60 velocity measurements are shown. It was difficult to determine whether ADP altered the velocity ratio because 150 μM ADP decreased microtubule velocity only slightly (~15%), whereas higher concentrations did not permit translocation by single motors.

‡ Two findings indicate that ATP- γS , and not a low level of contamination with ATP (0.5%), was responsible for the movement in these assays. First, the K_m for motility produced with ATP- γS was 8 μM (data not shown). At this concentration of total nucleotide, there is only 40 nM ATP, which by itself does not produce motility at detectable rates in these assays. Second, addition of 5 μM ATP (which on its own translocates microtubules at $0.025 \mu\text{m s}^{-1}$) to a solution of 250 μM ATP- γS did not change the velocity significantly.

§ In the single motor assay, translocation events lasting less than 0.5 μs or 0.5 s could not be distinguished from nonspecific adhesion to the surface or collisions with the surface caused by diffusion.

length before dissociating were assumed to be transported by individual motors (Fig. 1). Occasionally a microtubule dissociated from the motor before its trailing end reached the pivot point. By dividing the number of these events by the total translocation time for all microtubules, an average interaction time was determined.

The time spent by kinesin in its nucleotide-free state was prolonged by decreasing the ATP concentration to 5-fold below the K_m for ATPase activity (10 μM ; refs 7, 9). Under such conditions, microtubules moved by single kinesins dissociated

13-fold less frequently (once per 148 s; Table 1) than with 1 mM ATP (once per 11 s). The distance travelled per release was less than at high ATP concentrations, implying that dissociation occurs, albeit infrequently, during the nucleotide-free state. As seen previously⁵, single motors translocated microtubules at twice the velocity of multiple motors in a high-density assay, indicating that nucleotide-free kinesin binds strongly to microtubules and imposes a drag on movement.

ATP- γS is an analogue whose rate of hydrolysis, and possibly product release, is slow compared with that of ATP^{10,11}. In the presence of 250 μM ATP- γS , single motors translocated microtubules slowly ($0.012 \mu\text{m s}^{-1}$) and remained bound for a time (166 s) comparable to that observed with 2 μM ATP. The velocity ratio (single/multiple motors) of 1.4, which is between that of 2 μM ATP and 1 mM ATP ($P < 0.001$), suggests that ATP- γS prolongs a tightly bound microtubule state.

Two phosphate analogues, aluminium fluoride (250 μM) and vanadate (50 μM), were added to 1 mM ATP to prolong an intermediate similar to kinesin-ADP- P_i . In single and multiple motor assays with these analogues, some microtubules stopped translocating but remained bound, while others continued to move at close to full speed. Similar results have been found previously with vanadate in high-density assays¹². Immobile microtubules in the presence of aluminium fluoride rarely resumed movement. In the presence of vanadate, however, paused microtubules often resumed translocation (Fig. 2), perhaps because of an occasional dissociation of the inhibitor. Thus, both phosphate analogues cause strong binding of kinesin to microtubules. In contrast to the findings with 2 μM ATP and 250 μM ATP- γS , this strong binding state can resist the actions of other motors generating force on the same microtubule.

To extend a kinesin-ADP state, Mg-ADP was added as a competitive inhibitor to 1 mM ATP. The addition of 150 μM

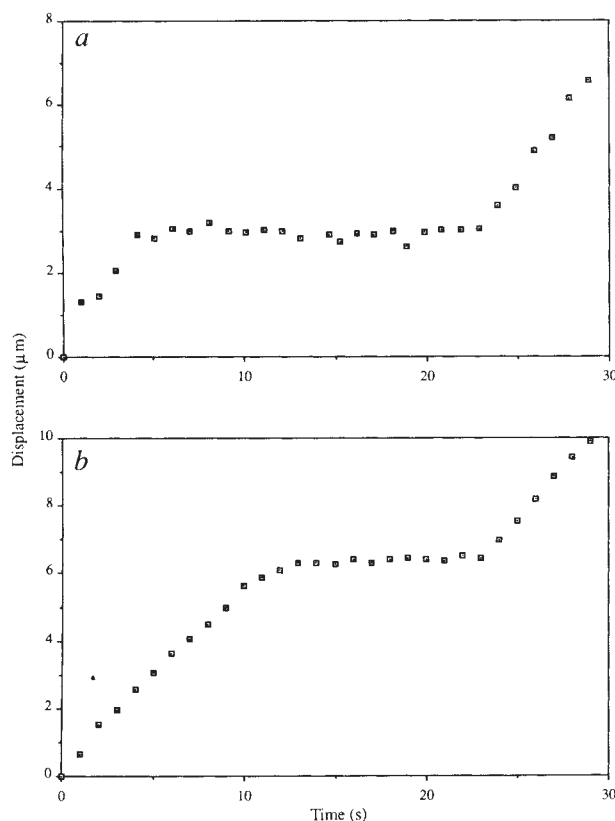


FIG. 2 Vanadate (50 μM) causes intermittent disruptions in microtubule transport in the presence of 1 mM ATP. The displacement of representative microtubules as a function of time at single (*a*) and multiple (*b*) motor densities is shown. The average velocities are 0.63 and $0.58 \mu\text{m s}^{-1}$ during translocation for *a* and *b* respectively. Paused microtubules resume translocation less frequently when either the motor density or the vanadate concentration is raised. But restarting occurred in almost all cases in the single motor assay with 50 μM vanadate. Microtubule ends ($\pm 0.2 \mu\text{m}$ error) were tracked using a custom computer program.

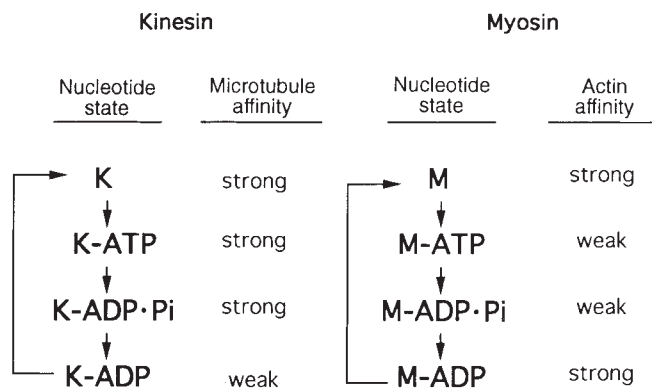


FIG. 3 Models for the filament binding interactions of kinesin (K) and myosin (M) during the ATP hydrolysis cycle. Possibilities of multiple nucleotide states or mechanical influence are not included in this scheme.

ADP caused microtubules to dissociate from single kinesins twice as frequently; no binding was detected with 250 μ M ADP (Table 1). Similar concentrations of ADP in the absence of ATP also dissociated microtubules from kinesin (data not shown). These results suggest that ADP weakens the binding between kinesin and microtubules. The sudden suppression of translocation with increasing ADP may reflect the need for both kinesin heads to detach simultaneously in order for a microtubule to dissociate. Consistent with this idea, higher concentrations of ADP (5–20 mM) were required to dissociate microtubules in the multiple motor assay.

These results suggest the following sequence of interactions between kinesin and microtubules during the ATP hydrolysis cycle (Fig. 3). In the nucleotide-free state, kinesin is strongly bound to microtubules. After nucleotide binding, the motor remains tightly associated with the polymer, yet some conformational change must occur that diminishes its ability to resist microtubule translocation in a multiple motor assay. Kinesin-ADP·Pi is still strongly bound to the microtubule, but after phosphate release, the interaction becomes weak, potentially allowing the motor to dissociate and find a new binding site. This model of the kinesin cycle requires the assumption that the analogues and inhibitors used here prolong the predicted states and that these states mimic normal intermediates in hydrolysis. Transient-state kinetic measurements^{13,14} of kinesin dissociation during the normal ATPase cycle will be needed to confirm this model.

In a motor protein's chemomechanical cycle, energy is required to dissociate the strongly bound motor-polymer complex and to produce force. Our results suggest that the utilization of ATP energy by kinesin is different from that of other cytoskeletal motors. Upon binding ATP, myosin¹³ and dynein^{14,15} rapidly dissociate from their filaments. Although the binding of ATP to kinesin is associated with a free energy change¹⁶ of similar magnitude to that of other motors^{17,18}, our results indicate that kinesin does not dissociate from microtubules early in its hydrolysis cycle. How then does kinesin use the binding energy of ATP? One possibility is that the energy may be stored in the protein for use later in the cycle, as occurs for other proteins such as bacteriorhodopsin¹⁹. Alternatively, the binding of ATP may be directly coupled to mechanical work. This issue can be resolved by determining when force production occurs in the hydrolysis cycle. □

- Johnson, K. A. *Rev. Biophys. biophys. Chem.* **14**, 161–188 (1985).
- Howard, J., Hudspeth, A. J. & Vale, R. D. *Nature* **342**, 154–158 (1989).
- Block, S. M., Goldstein, L. S. & Schnapp, B. J. *Nature* **348**, 348–352 (1990).
- Hackney, D. D. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6314–6318 (1988).
- Pate, E. & Cooke, R. *J. Musc. Res. Cell Motil.* **12**, 376–393 (1991).
- Kuznetsov, S. A. & Gelfand, V. I. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8350–8354 (1986).
- Goody, R. S. & Mannherz, H. G. in *Protein-Ligand Interactions* (eds Sund, H. & Blauer, G.) (Walter de Gruyter, Berlin, 1975).
- Shimizu, T., Katsura, T., Domanico, P. L., Marchese-Ragona, S. P. & Johnson, K. A. *Biochemistry* **28**, 7022–7027 (1989).
- Cohn, S. A., Ingold, A. L. & Scholey, J. M. *J. biol. Chem.* **264**, 4290–4297 (1989).
- Lynn, R. W. & Taylor, E. W. *Biochemistry* **10**, 4617–4624 (1971).
- Porter, M. E. & Johnson, K. A. *J. biol. Chem.* **258**, 6582–6587 (1983).
- Omoto, C. K. & Johnson, K. A. *Biochemistry* **25**, 419–427 (1986).
- Sadhu, A. & Taylor, E. W. *J. biol. Chem.* **267**, 11352–11359 (1992).
- Taylor, E. W. in *Heart and Cardiovascular System* (eds Fozzard, H. A.) 1281–1293 (Raven, New York, 1992).
- Holzbaur, E. L. F. & Johnson, J. A. *Biochemistry* **25**, 428–434 (1986).
- Mathies, R. A., Lin, S. W., Ames, J. B. & Pollard, W. T. *Rev. Biophys. Chem.* **20**, 491–518 (1991).
- Vale, R. D., Reese, T. S. & Sheetz, M. P. *Cell* **42**, 39–50 (1985).

ACKNOWLEDGEMENTS. We thank F. Malik for purifying kinesin, T. Shimizu and T. Mitchison for advice on purification of ATP- γ S, W. Marshall for developing the measurement program, R. Cooke for discussion, and F. Malik, G. McCaffrey and F. McNally for critically reading the manuscript. This work was supported by a grant from the NIH.

A role for reverse transcripts in gene conversion

Leslie K. Derr & Jeffrey N. Strathern

Laboratory of Eukaryotic Gene Expression, NCI Frederick Cancer Research and Development Center, Frederick, Maryland 21702, USA

RECOMBINATION between a diffusible reverse transcript and its homologous chromosomal allele has been proposed as a mechanism for the precise removal of introns from DNA and gene conversion of dispersed repeated sequences^{1,2}. We have reported that RNA-mediated recombination occurs in the yeast *Saccharomyces cerevisiae*³. This recombination requires expression of the retrotransposon Ty, and results in intron loss from a plasmid-borne marker gene and the formation of pseudogenes. Because the pseudogenes are embedded in Ty sequences, chromosomal insertion could have been mediated by Ty integrase or by homologous recombination with endogenous Ty sequences. The structure of the chromosomal recombinants and the fact that plasmid and chromosomal recombination can have different requirements demanded a direct demonstration of RNA-mediated gene conversion of a chromosomal allele. Here we report the first demonstration, to our knowledge, of recombination between a reverse transcript and its chromosomal homologue and describe an assay that specifically detects this novel recombination pathway.

To address the question of whether RNA-mediated recombination can contribute to mitotic gene conversion, we generated a system specifically to monitor chromosomal RNA-mediated events. This system has as its essential features a plasmid donor of complementary DNA and a homologous chromosomal target. The plasmid (Fig. 1a) carries a *his3* reporter gene whose coding sequence is interrupted by an artificial intron. The intron is inserted at a unique *MscI* site, in an unspliceable orientation relative to the *HIS3* promoter. Induction from a *GAL1* promoter fused to the 3' end of *his3* produces a spliceable transcript. Reverse transcription of this spliced antisense *his3* transcript can lead to *His3*⁺ prototrophy in a strain containing a chromosomal *his3* deletion (*his3-Δ200*)³. To provide a homologous chromosomal target, a 34-base-pair (bp) deletion spanning the *MscI* site of *his3*, *his3-ΔMscI*, was inserted at the *MAT* locus on chromosome III. DNA recombination between plasmid and chromosomal *his3* sequences cannot generate a *HIS3* allele. *His3*⁺ prototrophs can result from recombination between the *HIS3* cDNA and chromosomal *his3-ΔMscI* allele (Fig. 1d).

Because the presence of a homologous chromosomal target might increase the rate of formation of chromosomal *His3*⁺

Received 28 September; accepted 2 November 1992.

- Eisenberg, E. & Hill, T. L. *Science* **227**, 999–1006 (1985).
- Jencks, W. P. in *The Roots of Modern Biochemistry* 571–580 (Walter de Gruyter, Berlin, 1988).
- Cooke, R. *CRC crit. Rev. Biochem.* **21**, 53–118 (1986).

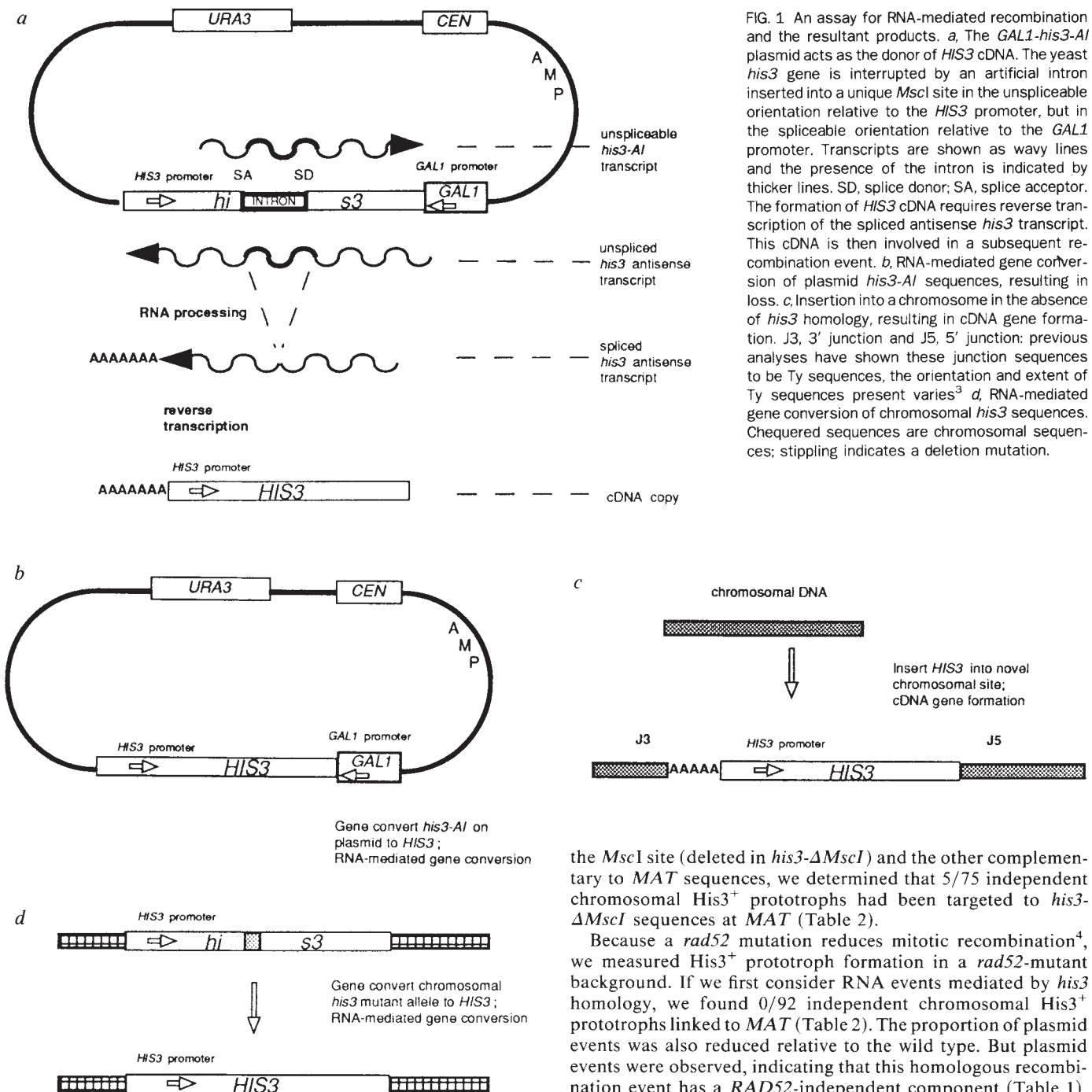


FIG. 1 An assay for RNA-mediated recombination and the resultant products. **a**, The *GAL1-his3-AI* plasmid acts as the donor of *HIS3* cDNA. The yeast *his3* gene is interrupted by an artificial intron inserted into a unique *MscI* site in the unspliceable orientation relative to the *HIS3* promoter, but in the spliceable orientation relative to the *GAL1* promoter. Transcripts are shown as wavy lines and the presence of the intron is indicated by thicker lines. SD, splice donor; SA, splice acceptor. The formation of *HIS3* cDNA requires reverse transcription of the spliced antisense *his3* transcript. This cDNA is then involved in a subsequent recombination event. **b**, RNA-mediated gene conversion of plasmid *his3-AI* sequences, resulting in loss. **c**, Insertion into a chromosome in the absence of *his3* homology, resulting in cDNA gene formation. J3, 3' junction and J5, 5' junction: previous analyses have shown these junction sequences to be Ty sequences, the orientation and extent of Ty sequences present varies³. **d**, RNA-mediated gene conversion of chromosomal *his3* sequences. Chequered sequences are chromosomal sequences; stippling indicates a deletion mutation.

recombinants, we compared the rate of His3⁺ prototroph formation in a strain lacking chromosomal *his3* homology (*his3-Δ200*) with a strain containing *his3-ΔMscI* target sequences (*his3-Δ200*, *MAT::his3-ΔMscI*). As seen in Table 1 (compare L1890 with YLD125), the rate of His3⁺ prototroph formation was not increased by the presence of chromosomal *his3* homology, nor was the proportion of events that were chromosomal increased by providing a *his3* target. Therefore, the rate of formation of chromosomal His3⁺ prototrophs is not limited by *his3* homology.

We then asked how often chromosomal His3⁺ prototrophs are the result of recombination with *his3-ΔMscI* target sequences. The most direct test of this was to assay for the presence of a 1-kilobase (kb) polymerase chain reaction (PCR) product that could only result if *HIS3* were linked to *MAT* (Fig. 2). Using one primer complementary to *HIS3* sequences spanning

the *MscI* site (deleted in *his3-ΔMscI*) and the other complementary to *MAT* sequences, we determined that 5/75 independent chromosomal His3⁺ prototrophs had been targeted to *his3-ΔMscI* sequences at *MAT* (Table 2).

Because a *rad52* mutation reduces mitotic recombination⁴, we measured His3⁺ prototroph formation in a *rad52*-mutant background. If we first consider RNA events mediated by *his3* homology, we found 0/92 independent chromosomal His3⁺ prototrophs linked to *MAT* (Table 2). The proportion of plasmid events was also reduced relative to the wild type. But plasmid events were observed, indicating that this homologous recombination event has a *RAD52*-independent component (Table 1). Overall, the rate of His3⁺ prototroph formation was not affected by the *rad52* mutation (Table 1). These results suggest that a *rad52*-independent mechanism of chromosomal insertion also exists. The simplest explanation, based on the structure of the chromosomal His3⁺ prototrophs, is that insertion is mediated by Ty integrase³ (see above).

The experiments described above suggested that at least three pathways existed for the insertion of *HIS3* cDNA: insertion by Ty integrase, *RAD52*-dependent gene conversion (plasmid and chromosomal events) and *rad52*-independent gene conversion (plasmid events only). To eliminate plasmid events and those events mediated by Ty integrase, the *HIS3* promoter and the initiation codon of the plasmid *his3-AI* sequences were deleted, generating a plasmid-borne *his3-ΔATG* allele. No (0/1.5 × 10¹⁰) His3⁺ events were observed in a *his3-Δ200* strain. But His3⁺ prototrophs were observed at a rate of 2.0 × 10⁻⁹ when RNA-mediated recombination with chromosomal *his3-ΔMscI* sequences was possible (Table 2). This 10-fold decrease in the rate of

TABLE 1 Rate of His³ prototroph formation

Strain chromosomal <i>his3</i> allele(s)	Rate ($\times 10^{-8}$)	<i>RAD52</i>	Chromosomal (%)
L1890 <i>his3-Δ200</i>	3.1	<i>RAD52</i>	53 (39/73)
YLD125 <i>his3-Δ200, his3-ΔMscI</i>	1.8	<i>RAD52</i>	53 (40/75)
YLD117 <i>his3-Δ200</i>	3.2	<i>rad52</i>	77 (57/74)
YLD157 <i>his3-Δ200, his3-ΔMscI</i>	3.9	<i>rad52</i>	75 (64/86)

Genotype of strains: L1890 (*MATa his3-Δ200 ura3-52 trp1-289 lys*); YLD125 (*MATa::his3-ΔMscI his3-Δ200 ura3-52 trp1-289 lys*); YLD117 (*MATa his3-Δ200 ura3-52 trp1-289 lys rad52-542*); YLD157 (*MATa::his3-ΔMscI his3-Δ200 ura3-52 trp1-289 lys rad52-542*). These isogenic strains carry the *URA3*-marked, CEN-based, *GAL1-his3-AI* plasmid. Rates were computed by the method of Lea and Coulson⁹ from the median tube in 75 (*RAD52*) and 100 (*rad52-542*) tube experiments. Isogenic strains carrying the *GAL1-his3-AI* plasmid were grown to saturation in media lacking uracil and containing galactose at 20 °C, and then scored on plates lacking histidine, with incubation at 30 °C. One His³ event was picked randomly from each culture and chromosomal events were distinguished from plasmid events by selecting for loss of the plasmid (*ura⁻*) by plating on FOA (ref. 10). Chromosomal events have a *ura⁻* His³ phenotype. The *rad52-542* allele was made by the method of Alani *et al.*¹¹. The *rad52-542* mutation was confirmed by MMS⁵ and Southern analysis (data not shown). The proportion of events that are chromosomal depending on the allele state of *RAD52* is significant at $\chi^2 = 6.11$, $P < 0.025$.

His³ prototroph formation corresponds nicely with the 3.5% ($0.53 \times 5/75$) His³ prototrophs linked to *MAT*. As expected, all of the His³ prototrophs were gene convertants of the *his3-ΔMscI* allele. This assay provides a means of assaying the cellular proteins specifically involved in recombination between a cDNA and its chromosomal allele.

RNA-mediated recombination offers a distinct pathway for gene conversion of chromosomal alleles. Spontaneous mitotic recombination between *his3* heteroalleles occurs at a 100-fold

TABLE 2 RNA-mediated gene conversion of chromosomal *his3-ΔMscI* sequences

Strain (<i>RAD52</i>)	Plasmid	Rate ($\times 10^{-8}$)	Chromosomal His ³ prototrophs linked to <i>MAT</i>
YLD125 (<i>RAD52</i>)	<i>GAL1-his3-AI</i>	1.8	5/75
YLD157 (<i>rad52</i>)	<i>GAL1-his3-AI</i>	3.9	0/92
YLD125* (<i>RAD52</i>)	<i>GAL1-his3-ΔATG</i>	0.2	52/52

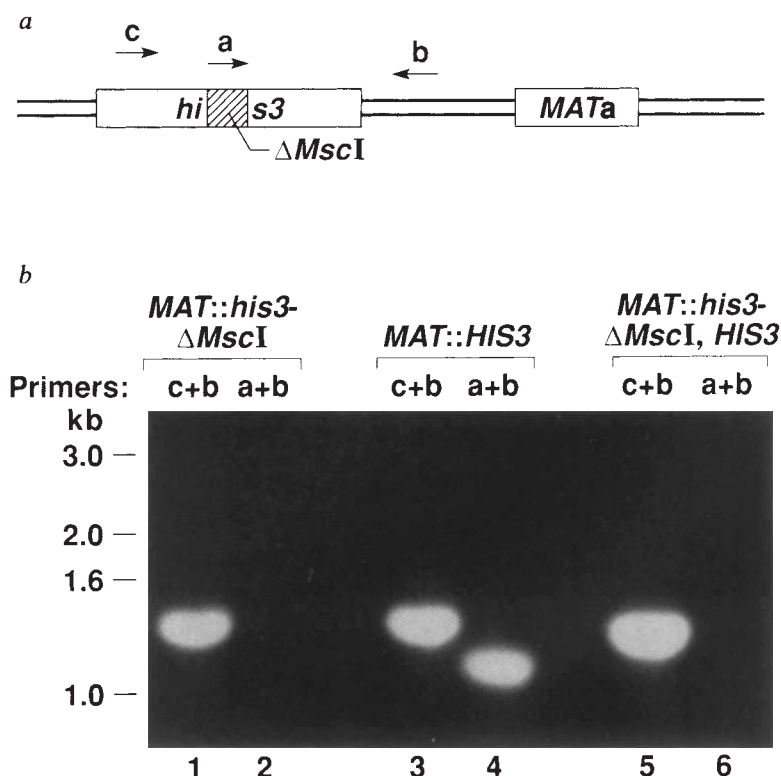
Strains are as in Table 1. Rates were computed as in Table 1. Chromosomal events were distinguished as in Table 1 and one chromosomal event was then picked from each culture. Independent (except for *GAL1-his3-ΔATG* see footnote below) chromosomal His³ prototrophs were analysed to determine the linkage of *HIS3* and *MAT*. His³ events linked to *MAT* were identified by the presence of a 1-kb PCR product and genetic linkage analysis. The observation that no His³ prototrophs are linked to *MAT* in *rad52-542* background is significant at $\chi^2 = 5.84$, $P < 0.025$.

* The *his3-ΔATG* mutation was created using PCR, resulting in the deletion of *HIS3* nucleotides -172 → +27 (ref. 12). Nine cultures, inoculated with 10^4 cells containing the *GAL1-his3-ΔATG* plasmid were grown to saturation in media lacking uracil and containing galactose at 20 °C, and then scored on plates lacking histidine, with incubation at 30 °C. The 52 events were pooled from the 9 independent cultures.

higher rate than RNA-mediated gene conversion⁵. To compare these RNA and 'DNA' rates, one should note that the alleles are different, but more importantly that the transcripts that can serve as potential donors are different. If the spontaneous events use an RNA-mediated pathway, then they act on the sense strand as opposed to the antisense *his3* strand. Additionally, the level of *his3* expression is different, the transcript produced here is from the efficient *GAL1* promoter, not the *HIS3* promoter. Any feature affecting the transcript, such as its abundance and its ability to serve as a substrate for reverse transcription, necessarily influences the rate of RNA-mediated recombination.

FIG. 2 A PCR-based assay to distinguish RNA-mediated gene conversions of homologous target sequences (*his3-ΔMscI*) from *HIS3* sequences inserted elsewhere. *a*, Structure of *his3-ΔMscI* on chromosome III and location of primers used for PCR. A 34-bp deletion (nucleotides 219–252; ref. 12) of the *his3* gene was constructed and inserted at *MAT*. This deletion spans the *MscI* site, the site of insertion of the artificial intron in plasmid *his3* sequences. *b*, Determination of the linkage between *HIS3* and *MAT*. The primers used for PCR are indicated above each lane. Lanes 1 and 2 show the starting parent strain YLD125. As expected no band is present with primers *a* and *b*. Lanes 3 and 4 show a His³ prototroph linked to *MAT*. Lanes 5 and 6 show a chromosomal His³ prototroph, but the *HIS3* sequences have been inserted elsewhere in a region lacking *HIS3* homology.

METHODS. Linkage of *HIS3* and *MAT* was determined by PCR, following loss of the *GAL1-his3-AI* plasmid. Primer *a*, 5'-CATGCTCTGGCCAAGCATTCC (nucleotides 222–242, ref. 12) is complementary to sequences spanning the *MscI* site, deleted in *his3-ΔMscI* and interrupted in plasmid *his3-AI* sequences. Primer *b* is complementary to *MAT* sequences 5'-CTGGGTAGAGTCTTATTGGCA (nucleotides 197379–197399; ref. 13). If *HIS3* is linked to *MAT*, then a 1-kb PCR product is expected. Primer *c*, 5'-CCCCGGCCGAATTCAGAGCA-GAAAGCCCTAGTA (nucleotides 27–47; ref. 12) was used in combination with primer *b* as a positive control. The conditions for PCR were: 3 μl yeast DNA, 300 ng each primer, 200 μM dNTPs, and 2.5 units *Taq* polymerase (Perkin Elmer Cetus); with denaturation at 94 °C for 1 min, annealing at 55 °C for 1 min and extension at 72 °C for 2 min, for 25 cycles, followed by a 7 min extension at 72 °C. The validity of this PCR measurement was confirmed by Southern and genetic linkage analyses.



Homologous insertion of cDNAs has two postulated effects: the removal of introns and the conversion of members of gene families^{1,2}. For this process to have evolutionary significance, the RNA-mediated gene conversion event must occur in germ line cells. Participation of RNA in this evolutionary event requires that it be expressed and accessible for reverse transcription. Retroviruses can provide a vehicle for horizontal transmission of non-germline expressed transcripts. It has been shown that nonretroviral RNA is packaged, reverse-transcribed and integrated into the host genome^{6,7}. Thus, RNA-mediated gene conversion could be important in germ-line substitutions, contributing to genome evolution, as well as somatic events. Kupiec has reported gene conversion of chromosomal Ty elements by Ty cDNA⁸. We have shown that gene conversion between a reverse transcript of a non-retroelement and its chromosomal allele occurs at an appreciable rate in *S. cerevisiae*. □

Received 14 September; accepted 12 November 1992.

1. Baltimore, D. *Cell* **40**, 481–482 (1985).
2. Fink, G. R. *Cell* **49**, 5–6 (1987).
3. Derr, L. K., Strathern, J. N. & Garfinkel, D. J. *Cell* **67**, 355–364 (1991).
4. Malone, R. E. et al. *Curr. Genet.* **14**, 211–223 (1988).
5. McGill, C. M. et al. *Curr. Genet.* **18**, 29–39 (1990).
6. Linial, M. *Cell* **49**, 93–102 (1987).
7. Dornburg, R. & Temin, H. M. *Molec. cell. Biol.* **8**, 2328–2334 (1988).
8. Melamed, C., Nevo, Y. & Kupiec, M. *Molec. cell. Biol.* **4**, 1613–1620 (1992).
9. Lea, D. E. & Coulson, C. A. *J. Genet.* **49**, 264–285 (1949).
10. Boeke, J. D., Lacroute, F. & Fink, G. R. *Molec. gen. Genet.* **197**, 345–346 (1984).
11. Alani, E., Cao, L. & Kleckner, N. *Genetics* **116**, 541–545 (1987).
12. Struhl, K. *Nucleic Acids Res.* **13**, 8587–8601 (1985).
13. Oliver, S. G. et al. *Nature* **357**, 38–46 (1992).

ACKNOWLEDGEMENTS. We thank B. Shafer for technical assistance and M. J. Curcio and S. L. Holbeck for critical evaluation of this manuscript. This work was supported in part by the National Cancer Institute, the DHHS with ABL and a Damon Runyon-Walter Winchell Cancer Research Fund Fellowship.

Arsenical-resistant trypanosomes lack an unusual adenosine transporter

Nicola S. Carter & Alan H. Fairlamb*

Department of Medical Parasitology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

THE melaminophenyl arsenical melarsoprol is still used to treat African sleeping sickness^{1–3}, a disease caused by parasitic protozoa of the *Trypanosoma brucei* subgroup. Based on the observation that melamine antagonizes the trypanocidal activity of this class of drugs⁴, we investigated whether other physiological compounds could compete for the same receptor. Here we report that the *in vitro* trypanolytic effect of melarsen oxide can be specifically abrogated by adenine, adenosine and dipyrindamole, all of which compete for uptake by an adenosine transporter. Melarsen-sensitive trypanosomes have two high-affinity adenosine transport systems: a P1 type, which also transports inosine; and a P2 type, which also transports adenine and the melaminophenyl arsenicals. Melarsen-resistant trypanosomes lack P2 adenosine transport, suggesting that resistance to these arsenicals is due to loss of uptake.

As radiolabelled drug was unavailable, a direct approach towards identifying the mode of uptake and the intracellular targets for these drugs was not possible. Instead we used a simple spectrophotometric assay^{5,6} to identify compounds that might antagonize the trypanolytic effect of arsenicals against bloodstream forms of *Trypanosoma brucei brucei*. Lysis by both melarsen oxide and phenylarsine oxide (which lacks the

TABLE 1 Effects of 1 mM inosine on adenosine transport in a melarsen sensitive and resistant *Trypanosoma brucei brucei* cloned line

Phenotype	Uptake (pmol/10 ⁸ cells/sec)		
	No inosine (P1 and P2)	+1 mM inosine (P2 only)	Net (residual P1)
Sensitive	10.7 ± 1.1 (6)	4.7 ± 0.5 (6)	6.0
Resistant	1.9 ± 0.3 (5)	ND (6)	1.9

Transport was measured with 1 μ M [³H]adenosine (final specific activity, 0.56 Ci mmol⁻¹) at 25 °C in the presence (P1-inhibited) or absence of 1 mM inosine as described for Fig. 3, in two cloned lines of *Trypanosoma brucei brucei*, the parent S427 c118 (ref. 32), which is susceptible to melaminophenyl arsenicals, and a derived melarsen-resistant line S427 cRU15 (ref. 6). Results are expressed as a mean value (number of determinations given in brackets) plus or minus standard error of the mean. ND, No detectable rate could be measured and regression coefficients were 0.67 or less.

melamine ring and has no activity *in vivo*⁷) are dependent on both concentration and time with rapid lysis occurring at concentrations greater than 0.5 μ M (Fig. 1a and b). Compounds were screened at a 100-fold excess (50 μ M) for their ability to block lysis. These included a selection of basic amino acids, polyamines, nitrogen-containing vitamins and coenzymes, and the commonly occurring purines and pyrimidines and their nucleotides. Of the 32 compounds tested, only the purines, adenine and adenosine, were able to significantly delay lysis by melarsen oxide in a dose-dependent manner (Fig. 1c and e). In sharp contrast, addition of these purines in up to a 100-fold excess had no effect on lysis induced by phenylarsine oxide (Fig. 1d and f). Dipyrindamole, a general nucleoside transport inhibitor^{8,9}, was also a potent inhibitor of melarsen oxide-induced lysis, but again had no effect on phenylarsine oxide-induced lysis (Fig. 1g and h).

As the Trypanosomatidae have no *de novo* synthesis of purines and instead salvage pre-formed purines from the host^{10,11} we investigated whether this receptor could be part of an adenine/adenosine carrier by measuring [³H]adenosine transport into the trypanosome. As uptake was extremely rapid, transport kinetics were measured at 1-second intervals over the first five seconds, and initial rates were determined by linear regression analysis. As detailed below, we can identify two distinct adenosine transporters in *T. brucei*, consistent with earlier observations in other Trypanosomatidae^{9,12,13}. For convenience we have named these transport systems P1 and P2 (Fig. 2). Uptake of 1 μ M adenosine via the P1-transporter can be inhibited in a dose-dependent and saturable manner by inosine, accounting for 60–70% of total adenosine uptake (Fig. 3a, open symbols). Likewise, uptake of adenosine via the P2 transporter can be inhibited by increasing concentrations of adenine, accounting for 30–40% of total adenosine uptake into the cell (Fig. 3b, open symbols). When saturating concentrations of both adenine and inosine are present, adenosine transport is completely inhibited (Fig. 3a and b, closed symbols).

By saturating one transporter with either 1 mM inosine (P1) or 100 μ M adenine (P2), the kinetic parameters for [³H]adenosine uptake of either transporter can be shown to obey simple Michaelis-Menten kinetics (Fig. 3c and d). The apparent K_m values for the P1 and P2 transporters are 0.15 ± 0.02 and 0.59 ± 0.05 μ M, respectively, with V_{max} values of 10.60 ± 0.53 and 9.5 ± 0.27 pmol per 10⁸ cells per second, respectively. As expected, adenine competitively inhibits adenosine transport on the P2 transporter with a K_i of 0.38 μ M (Fig. 3e). Moreover, the P2, but not the P1 transporter can be markedly inhibited by 1 μ M melarsen oxide or melarsoprol (not shown). In the case of melarsoprol, kinetic analysis shows competitive-type inhibition with a K_i of 0.28 μ M (Fig. 3f). Nucleoside transporters in other Trypanosomatidae^{14–16} and in mammalian cell types⁸ usually have a broad substrate specificity, transporting both

* To whom correspondence should be addressed.

purines and pyrimidines. This is not the case with the P2 transporter which is unable to transport other purines and pyrimidines (not shown). This may explain why only African trypanosomes of the *brucei* subgroup are susceptible to chemotherapy with melaminophenyl arsenicals⁷.

Drug resistance to another type of arsenical, trypanamide, has been shown to be associated with a marked decrease in drug uptake¹⁷. If melaminophenyl arsenical uptake into the trypanosome is mediated by an adenine/adenosine transporter, then a loss or conformational change in the P2 transporter could also

lead to arsenical resistance. We investigated this possibility using a cloned melarsen-resistant line derived *in vivo* from the susceptible parent clone⁶. The melarsen-resistant clone shows marked cross-resistance *in vivo* to all of the melaminophenyl arsenicals including melarsoprol (67-fold resistant) and melarsen oxide (33-fold resistant), but not to phenylarsine oxide *in vitro*⁶. The results shown in Table 1 indicate that adenosine transport in the resistant clone is markedly different: P2 adenosine transport is absent and the residual transport is threefold less than P1 transport in the susceptible clone. The lack of P2 adenosine

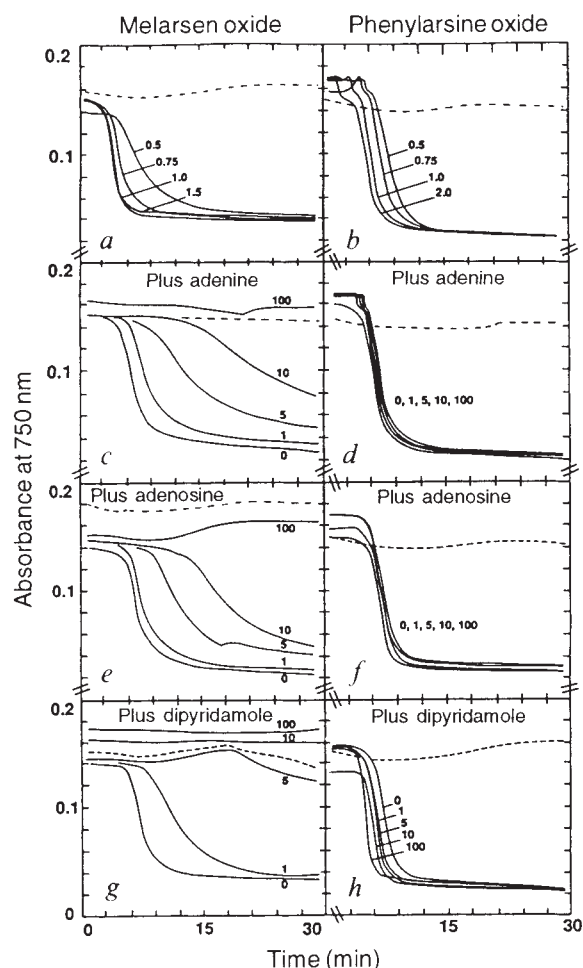


FIG. 1 Effect of adenine, adenosine and dipyrindamole on arsenical-induced lysis in *T. brucei*. Broken line represents control incubations containing no drug. *a*, Melarsen oxide-induced lysis measured at 0.5 to 1.5 μM . *b*, Phenylarsine oxide-induced lysis measured at 0.5 to 2.0 μM . *c*, Protection by adenine of melarsen oxide-induced lysis; melarsen oxide 0.5 μM plus or minus 1, 5, 10 and 100-fold excess adenine. *d*, Protection by adenine of phenylarsine oxide-induced lysis; phenylarsine oxide 0.5 μM plus or minus 1, 5, 10 and 100-fold excess adenine. *e*, Protection by adenosine of melarsen oxide-induced lysis; conditions as for adenine in *c*. *f*, Lack of protection by adenosine of phenylarsine oxide-induced lysis; conditions as for adenine in *d*. *g*, Protection by dipyrindamole of melarsen oxide-induced lysis; conditions as for adenine in *c*. *h*, Lack of protection by dipyrindamole of phenylarsine oxide-induced lysis; conditions as for adenine in *d*.

METHODS. Bloodstream *Trypanosoma brucei brucei* (S427 c118)³² were isolated³³, and resuspended at 4 °C in an RPMI-derived basal salt solution, CBSS (ref. 6). Cell lysis was measured spectrophotometrically by monitoring the decrease in absorbance due to light scatter at 750 nm (refs. 5, 6). All solutions were pre-warmed to 37 °C. At zero time 1×10^7 cells were added to a cuvette containing CBSS and the required concentration of arsenical and test compound. Absorbance was followed for 30 min in a thermostat-controlled Beckman DU-70 singlebeam spectrophotometer fitted with an automatic 6-cell sample changer. Control samples were included in each assay. The arsenical drugs, melarsen oxide and melarsoprol were a gift from Rhône-Poulenc. Other chemicals were analytical grade purchased from Sigma, BDH or Aldrich.

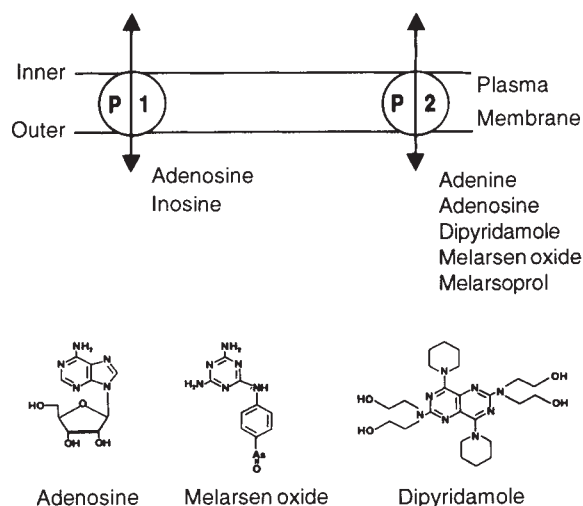


FIG. 2 An overview of adenosine transport in *T. brucei*. An illustration of the substrate specificities of the P1-type and P2-type transporters and a comparison of the chemical structures of melarsen oxide, adenosine and dipyrindamole, all of which are recognized by the P2 transporter.

transport in the melarsen-resistant trypanosomes suggests that they may also lack the ability to transport arsenicals, accounting for the observed cross-resistance patterns.⁶

Bloodstream trypanosomes lack a functional tricarboxylic acid cycle and are entirely dependent on substrate-level phosphorylation via glycolysis for ATP production¹⁸. Rapid loss of motility and cell lysis occurs following inhibition of glycolysis by salicylhydroxamic acid plus glycerol^{19,20}, as observed here

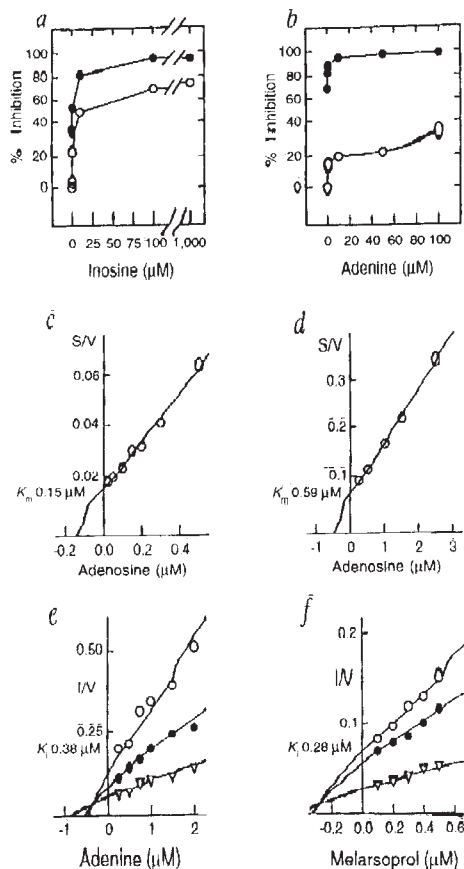


FIG. 3 Characterization of adenosine transport (at 25 °C) in *T. brucei*. *a*, Inhibition of $[^3\text{H}]$ adenosine transport by varying concentrations of inosine in the absence (○) and presence (●) of 100 μM adenine. *b*, Inhibition of $[^3\text{H}]$ adenosine transport by varying concentrations of adenine in the absence (○) and presence (●) of 1 mM inosine. *c*, Kinetics of uptake of $[^3\text{H}]$ adenosine via the P2 transporter in the presence of 100 μM adenine to saturate the P1 transporter. *d*, Kinetics of uptake of $[^3\text{H}]$ adenosine via the P2 transporter in the presence of 1 mM inosine to saturate the P1 transporter. *e*, Dixon plot showing competitive inhibition of uptake of $[^3\text{H}]$ adenosine by adenine via the P2 transporter. Transport of $[^3\text{H}]$ adenosine via the P1 transporter was competitively blocked by 1 mM inosine. Concentrations of $[^3\text{H}]$ adenosine were fixed at 0.5 (○), 1.0 (●) and 2.5 μM (▽), and the inhibitor, adenine, varied between 0.25 to 2.0 μM . *f*, Dixon plot showing competitive inhibition of uptake of $[^3\text{H}]$ adenosine by melarsoprol via the P2 transporter. Conditions as in *e*, only substituting melarsoprol (0.1–0.5 μM) for adenine.

METHODS. Purified bloodstream *T. brucei*, S427 c118 (ref. 33), were resuspended at 4×10^8 cells ml^{-1} in CBSS (ref. 6) at 4 °C. Uptake was measured at 1 s intervals using an adaptation of a previously described rapid sampling technique³⁴. CBSS (100 μl at 25 °C) containing radiolabelled $[2,5\text{-}^3\text{H}]$ adenosine (56 Ci mmol^{-1} ; Amersham) and unlabelled adenosine and inhibitors at 2 $\times$ final concentration, was overlaid on 100 μl of silicone oil (75 centistokes, 1.05 g ml^{-1} ; from Medford Silicones, NJ) in 1.5 ml Eppendorf tubes. At zero time, 100- μl aliquots of cells (pre-warmed to 25 °C) were pipetted sequentially at 1-s intervals into each of the tubes. Transport was stopped in each sample by pelleting the cells through the oil layer in an Eppendorf microfuge. Samples were prepared for liquid Scintillation counting³⁸ and radioactivity determined with 1 ml Picofluor 40 scintillation fluid (Canberra Packard) in a Beckman LS 6000LL (Beckman Instruments). The initial rate of uptake was determined by linear regression analysis (only regression coefficients greater than 0.95 were used). The results in *c* and *d* were analysed using the Enzfitter software package (Elsevier, Biosoft).

with trivalent arsenicals. Consequently, the mode of action of arsenicals has been attributed to primary inhibition of glycolysis, specifically trypanosomal pyruvate kinase²¹. More recently, the decrease in glycolytic flux was shown to be secondary to cell lysis and to not involve inhibition of pyruvate kinase²². Although the mode of action of arsenicals is unclear, one primary event is the formation of a stable adduct (MEL T)²³ between melarsen oxide and dihydrotrypanothione²⁴, a unique intracellular dithiol. However, levels of dihydrotrypanothione and glutathione are not significantly different in the melarsen-sensitive and resistant cloned lines described here⁶. Although the lipoic acid content is reduced twofold in the resistant clone²⁵, melarsen oxide, free or complexed to dihydrolipoamide, does not potently inhibit the plasma membrane-associated dihydrolipoamide dehydrogenase²⁶. In *Escherichia coli*, lipoic acid has been implicated in the transport of ribose, galactose and maltose²⁷. Whether lipoic acid is involved in the uptake of arsenical drugs in trypanosomes or serves as a drug target has not been established. Nonetheless, our current results suggest that an important feature of the selective toxicity of these drugs against trypanosomes is conferred by their differential uptake by the parasite.

The unusual nature of purine transport in African trypanosomes could be exploited for future chemotherapy. First, the apparent lack of P2 adenosine transport might make melarsen-resistant trypanosomes hypersensitive to attack by purine analogues recognized by other purine transporters. However, allopurinol, a hypoxanthine analogue with therapeutic activity against other members of the Kinetoplastida^{28,29}, was completely ineffective *in vivo* against both melarsen-sensitive and resistant *T. brucei*⁶. Moreover, both cloned lines were equally sensitive to the inosine analogue Formycin B (ref. 6). Second, the structural similarities between adenine and melamine suggests that it is the melamine moiety which specifically targets these drugs into the parasite via the P2-transporter (Fig. 2). This supports Paul Ehrlich's original postulate (in 1898) that drugs comprise a 'haptophoric' group that specifies cellular recognition and a 'toxophoric' group that exerts the specific pharmacological effect³⁰. The unusual specificity of this P2 transporter might be useful for the selective delivery of 'toxophores' other than arsenic into melarsen-susceptible trypanosomes, thus avoiding the often lethal side effect of arsenical-induced encephalopathy in the host^{2,3,31}. □

Received 11 September; accepted 3 December 1992.

- Friedheim, E. A. H. *Am. J. trop. Med. Hyg.* **29**, 173–180 (1949).
- Apted, F. I. C. *Pharmac. Ther.* **11**, 391–413 (1980).
- Wery, M. *Curr. Opin. Infect. Dis.* **4**, 838–843 (1991).
- Williamson, J. Br. *J. Pharmac.* **14**, 431–442 (1959).
- Mesnick, S. R., Blobstein, S. H., Grady, R. W. & Cerami, A. *J. exp. Med.* **148**, 569–579 (1978).
- Fairlamb, A. H., Carter, N. S., Cunningham, M. & Smith, K. *Molec. biochem. Parasit.* **53**, 213–222 (1992).
- Hawking, F. in *Experimental Chemotherapy* vol. 1 (eds Schnitzer, R. J. & Hawking, F.) 129–256 (Academic, New York, 1966).
- Plagemann, P. G. W., Wohliuter, R. M. & Woffendin, C. *Biochem. biophys. Acta* **947**, 405–443 (1988).
- James, D. M. & Born, G. V. R. *Parasitology* **81**, 383–393 (1980).
- Hammond, D. J. & Gutteridge, W. E. *Molec. biochem. Parasit.* **13**, 243–261 (1984).
- Fairlamb, A. H. *Parasitology* **99S**, 93–112 (1989).
- Okochi, V. I., Abaelu, A. M. & Akirimi, E. O. *Biochem. Int.* **6**, 129–139 (1983).
- Baer, H. P., Serigne, V., Ogunjode, P. O. J. & Dzimir, M. *Am. J. trop. Med. Hyg.* **47**, 87–91 (1992).
- Ogunjode, P. O. J., Al-Jaser, M. H. & Baer, H. P. *Exp. Parasit.* **73**, 369–375 (1991).
- Finley, R. W., Cooney, D. A. & Dvorak, J. A. *Molec. biochem. Parasit.* **31**, 133–140 (1988).
- Aronow, B., Kaur, K., McCartan, K. & Ullman, B. *Molec. biochem. Parasit.* **22**, 29–37 (1987).
- Hawking, F. *Ann. trop. Med. Parasit.* **32**, 313–331 (1938).
- Fairlamb, A. H. & Oppendoes, F. R. in *Carbohydrate Metabolism in Cultured Cells* (ed. Morgan, M. J.) 183–224 (Plenum, New York, 1986).
- Clarkson, A. B. Jr & Brohn, F. H. *Science* **194**, 204–206 (1976).
- Fairlamb, A. H., Oppendoes, F. R. & Borst, P. *Nature* **265**, 270–271 (1977).
- Flynn, I. W. & Bowman, I. B. R. *Comp. biochem. Physiol.* **48B**, 261–273 (1974).
- Van Schaftingen, E., Oppendoes, F. R. & Hers, H. *Eur. J. biochem.* **166**, 653–661 (1987).
- Fairlamb, A. H., Henderson, G. B. & Cerami, A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2607–2611 (1989).
- Fairlamb, A. H., Blackburn, P., Ulrich, P., Chait, B. T. & Cerami, A. *Science* **227**, 1485–1487 (1985).
- Fairlamb, A. H., Smith, K. & Hunter, K. J. *Molec. biochem. Parasit.* **53**, 223–232 (1992).
- Danson, M. J., Conroy, K., McQuattie, A. & Stevenson, K. J. *Biochem. J.* **243**, 661–665 (1987).
- Richarme, G. *J. Bact.* **162**, 286–293 (1985).
- Nelson, D. J. et al. *J. Biol. Chem.* **254**, 11544–11549 (1979).
- Marr, J. J., Berens, R. L. & Nelson, D. J. *Science* **201**, 1018–1020 (1978).
- Ehrlich, P. in *The Collected Papers of Paul Ehrlich*, Vol. 1. *Histology, Biochemistry and Pathology* (ed. Himmelweit, F.) 596–618 (Pergamon, London, 1956).

31. Haller, L., Adams, H., Merouze, F. & Dago, A. *Am. J. trop. Med. Hyg.* **35**, 94–99 (1986).
 32. Hooijmakers, J. H. J., Borst, P., Van den Burg, J., Weissmann, C. & Cross, G. A. M. *Gene* **8**, 391–417 (1980).
 33. Lanham, S. M. *Nature* **218**, 1273–1274 (1968).
 34. Aronow, B., Allen, K., Patrick, J. & Ullman, B. *J. biol. Chem.* **260**, 6226–6233 (1985).

ACKNOWLEDGEMENTS. We thank P. Ulrich for the initial strain resistant to sodium melarsen. This work was supported by the UNDP/WORLD BANK/WHO Special Programme for Research and Training in Tropical Diseases. N.S.C. received support from the MRC and A.H.F. from the Wellcome Trust.

Ordered duplex RNA controls capsid architecture in an icosahedral animal virus

Andrew J. Fisher* & John E. Johnson†

Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907-1392, USA

* Present address: Institute for Enzyme Research, University of Wisconsin, Madison, Wisconsin 53705, USA

† To whom correspondence should be addressed

SMALL spherical viruses are among the simplest replicating systems in biology, yet the factors affecting their assembly, stability and disassembly are still poorly understood. A molecular switch is required for the assembly of icosahedral virus particles containing more than 60 identical subunits because strict symmetry cannot be maintained in subunit packing¹. All previously reported viruses with this type of structure use a portion of the capsid protein to regulate interactions between chemically equivalent but structurally distinct interfaces^{2–4}. We have investigated the $T = 3$ quasi-equivalent⁵ nodaviruses, which are small non-enveloped viruses with a single-stranded RNA genome that infect insects⁶, mice⁷ and fish⁸. They undergo a well-characterized series of steps in assembly and maturation^{9,10}, which in some respects are similar to the picornaviruses¹¹, despite their different capsid architecture. Here we report the X-ray structure of Flock House virus at 3.0 Å resolution, which reveals an ordered RNA duplex of 20 nucleotides and a protein segment that control the subunit interactions in this

TABLE 1 X-ray data collection for FHV

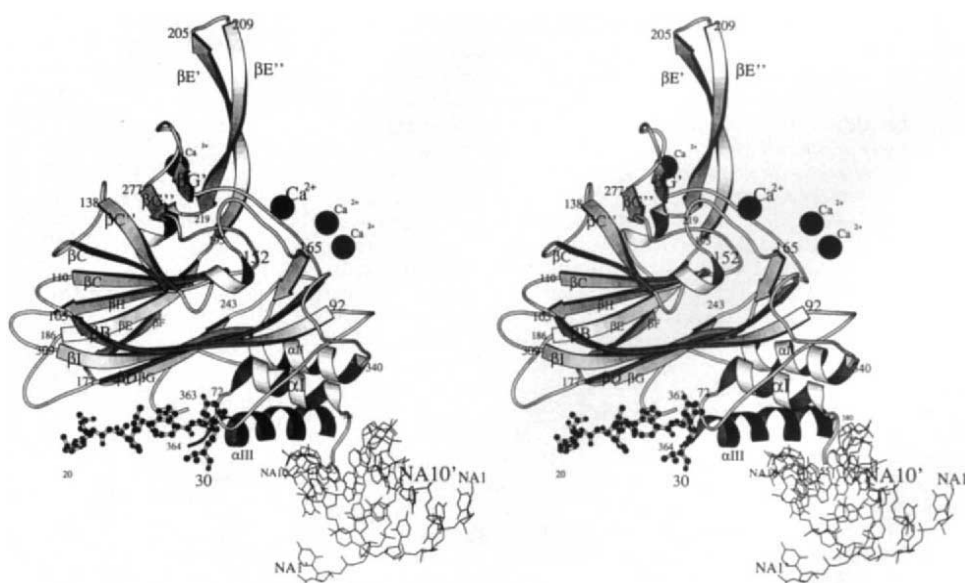
Resolution (Å)	Unique reflections	Percentage of data	Correlation*
50.0–15.0	3,753	74	ND
15.0–10.0	10,229	84	0.87
10.0–8.0	14,177	87	0.91
8.0–6.0	39,203	84	0.92
6.0–5.0	48,193	81	0.90
5.0–4.0	103,108	78	0.85
4.0–3.5	87,626	66	0.75
3.5–3.0	102,420	42	0.53
Overall	408,709	64	0.84
Number of films	167		
$R_{\text{merge}}^{\dagger}$	13.0		

Virus propagation, crystallization and preliminary data analysis, including orientation, have been previously described¹⁹. Diffraction data were collected at CHESS A1 line. A polyaniline model of black beetle virus²⁰ was positioned into the FHV R3 cell and initial phases computed for all the observed data to 3 Å resolution. Molecular replacement was started at 4.0 Å resolution because of the resemblance in structures predicted from the 87% identity in coat protein sequence between FHV and BBV. Phases were refined by averaging the electron density map over the 20-fold non-crystallographic symmetry followed by inverse Fourier transformation²¹. After 14 cycles of phase refinement, the overall correlation coefficient was 0.90. Resolution of the electron density map was then extended to 3.5 Å. The additional high-resolution data included in this calculation initially incorporated phases from the polyaniline model, which were further refined. The correlation coefficient was 0.89 after ten cycles of phase refinement at 3.5 Å resolution. Resolution was similarly extended to 3.25 Å then to 3.0 Å. The overall correlation coefficient for the last cycle was 0.84 for 15–3 Å data. The FHV model was built up by altering the BBV structure with the program FRODO²². The electron density map was readily interpretable and the FHV structure was built for all three subunits. The modelled C subunit consists of residues 20–30, 55–363 (of β), and 364–380 (of the cleavage product γ). Residues 1–19 and 31–54 contain many basic residues and probably interact with RNA that is not icosahedrally ordered. Residue 57 is the first observed amino acid for the A and B subunits, with residue 379 being the last ordered in the A subunit and 360 the last for the B subunit. ND, not determined.

* Correlation coefficient = $\sum (\langle F_o \rangle - |F_o|)(\langle F_c \rangle - |F_c|) / [\sum (\langle F_o \rangle - |F_o|)^2 \cdot (\langle F_c \rangle - |F_c|)^2]^{1/2}$.

† $R_{\text{merge}} = [\sum_h \sum_i (\langle I_h \rangle - I_{hi}) / \sum_h \sum_i I_{hi}] \times 100$, where $\langle I_h \rangle$ is the mean of the I_{hi} observations of reflection h .

FIG. 1 Stereo ribbon diagram showing the tertiary structure of the C subunit of FHV (see Fig. 2 for definition of C subunit). The subunit in this figure is oriented so that the virus exterior is at the top and the interior is at the bottom. The peptide arm (residues 20–30), which is only ordered in the C subunit, is represented as a ball-and-stick model. Residues 1–19 and 31–54, which are not visible in the electron density, contain many basic amino acids. The autocatalytic maturation cleavage occurs between Asp 363 and Ala 364, and the cleavage product (γ) is coloured black. Residues 381 to the carboxy-terminal 407 are disordered. The duplex RNA is illustrated as lines showing atomic bonds. NA1 is the 5' end and NA10 is the 3' end of the RNA; NA1' and NA10' label the 2-fold-related RNA strand. Calcium is known to stabilize the provirion structure of FHV⁹. The location of Ca^{2+} ions was identified from a difference electron density map calculated with coefficients $|F_{\text{native}}| - |F_{\text{EGTA}}| e^{-f_{\text{native}}}$, where $|F_{\text{EGTA}}|$ is the structure factor amplitude measured from crystals soaked in 40 mM EGTA. Five peaks, all coordinated by oxygen atoms, were evident in each icosahedral asymmetric unit (Fig. 2d). The two Ca^{2+} ions shown in the upper left are



related by quasi-3-fold symmetry and the two at the lower right are located on the quasi-3-fold axes. A similar observation was made for $T = 3$ plant viruses^{23,24}.

animal virus. The RNA interacts with a helical protein domain of the subunit that lies inside the capsid shell. One of the helices that binds the RNA is part of a 44-amino-acid polypeptide which is autocatalytically cleaved from the initial subunit translation product after virion assembly. The structure indicates that RNA associated with the cleaved polypeptide may be important in the infection process.

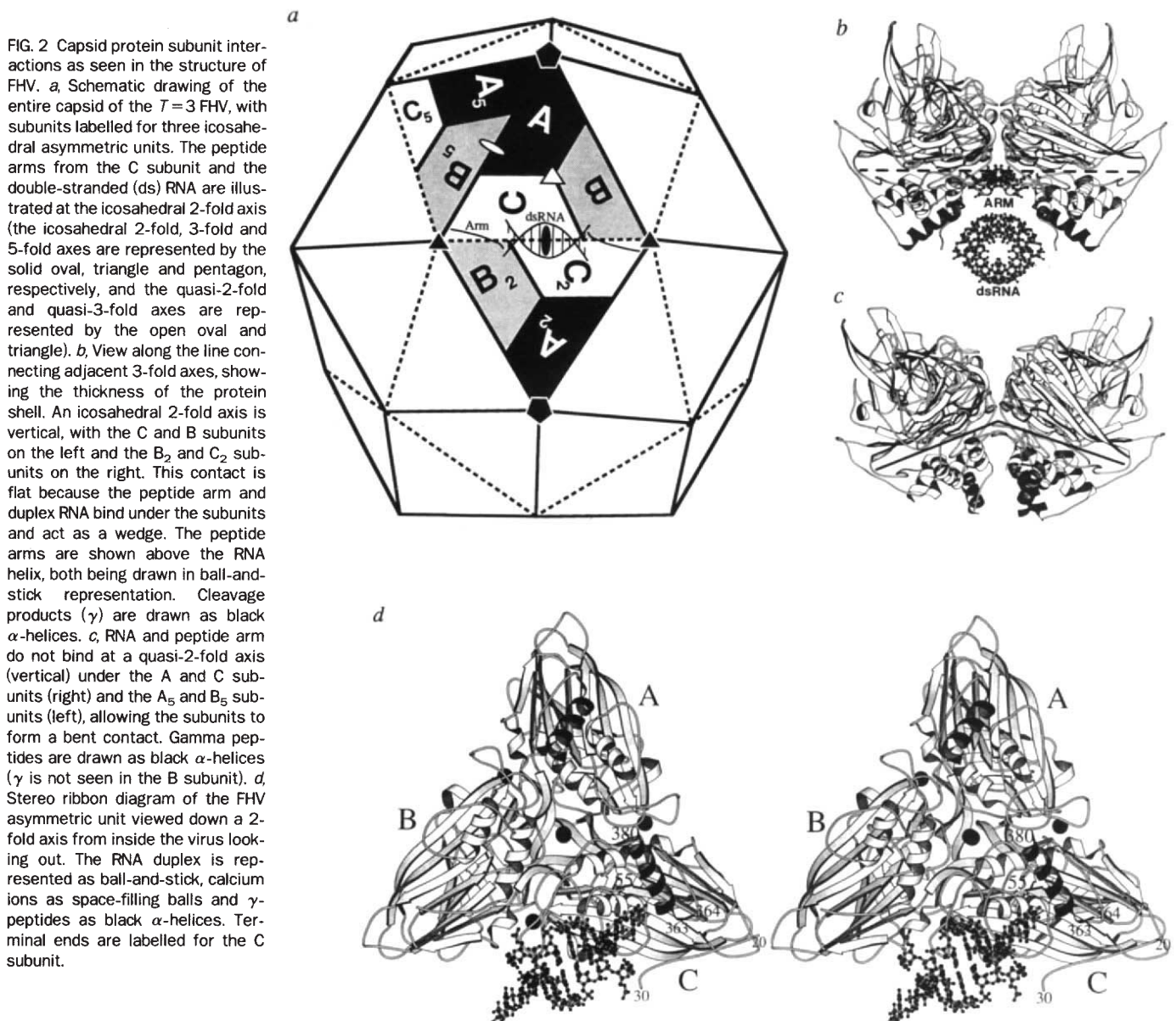
Table 1 summarizes the X-ray data collection and structure determination for Flock House virus (FHV). The subunit structure of FHV is very similar to that of black beetle virus (BBV) (Fig. 1). The viral capsid is composed of 60 triangular units, which consist of three identical polypeptides (designated A, B and C) related by icosahedral symmetry (Fig. 2*a*). All three subunits contain a central motif of an eight-stranded antiparallel β -barrel, similar to the topology of many other virus structures¹². The exterior surface consists of elaborate loops between the β -strands and the inner surface is made up of a helical domain from the amino- and carboxy-terminal ends.

Quasi-equivalence between the A, B and C subunits breaks down at residues 20–30, where they constitute an ordered arm in the C polypeptide at the flat joint separating the C and B₂ subunits (Figs 1 and 2). The packing of the A and B subunits at the quasi-2-fold axis is bent because residues 20–30 from the

A and B polypeptides and the RNA described below are not ordered at that interface.

Electron density comparable in height to protein but at a lower radius was readily modelled as duplex RNA (Fig. 3*a*). The RNA duplex lies on the icosahedral 2-fold axis, so one strand is related by 2-fold symmetry to the other. Ten nucleotides were modelled into the density for one strand (labelled NA1–NA10, 5' to 3', and related by 2-fold symmetry, NA1'–NA10'). The map displayed clear phosphate, sugar and base electron density for seven nucleotides, establishing the polarity of the strand. Three nucleotides at the 5' end were fitted into density that fixed the sugar-phosphate backbone, but was deficient in density corresponding to the bases. The four bases in the middle of each strand base-pair with their symmetry counterpart, and the three bases at either end maintain the helical nature of the duplex; but no base pairing is observed because of the absence of density for three bases at the 5' end in each strand. Sugars for all the nucleotides were modelled with the C3'-endo configuration, creating a phosphate distance between successive nucleotides of 6.0 Å, which agreed with the density.

The density for the bases must represent an average as there is not a repeat of 10 nucleotides that occurs 60 times in the genomic sequence. Pyrimidines could easily be fitted into the



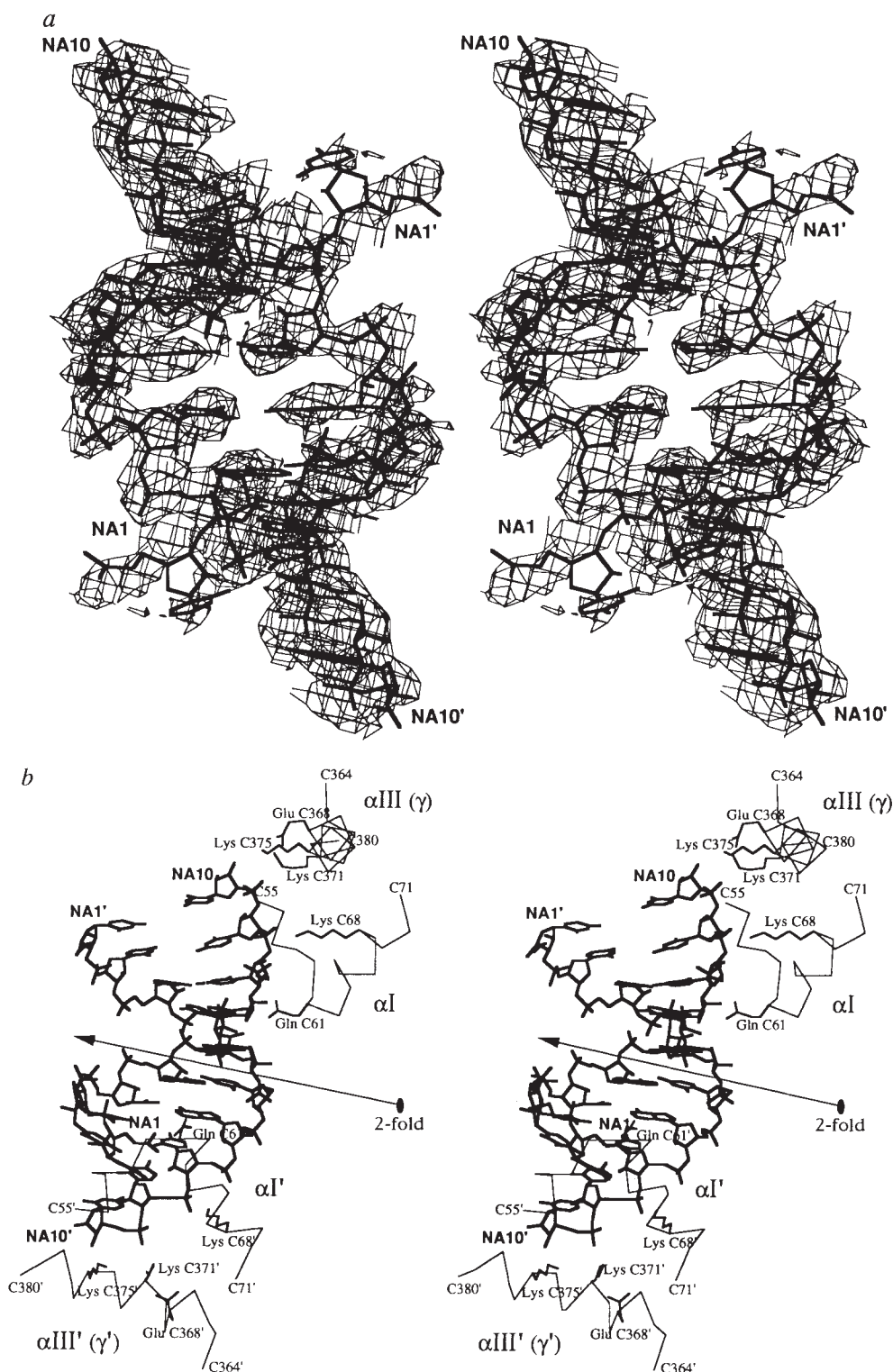
seven nucleotides with base density. But purines were modelled in positions NA5 and NA7 to maintain Watson-Crick base-pairing for the four middle bases, although the purines did not exactly fit the density. The duplex structure as a whole approximates that of an A-type helix. The helical pitch is 10.73 residues, with a 3.07 Å rise and a 33.55° helix rotation per residue and an average base-pair tilt of 8.3° for the four base pairs.

The RNA duplex binds in a groove under the C-C₂ subunit interface and interacts only with the internal helices α I and α III (γ) of the C subunit and its 2-fold-related neighbour (Fig. 3b). The contacts are electrostatic and van der Waals with the sugar-phosphate backbone. Lysines 68, 371 and 375 that salt-

link with the phosphates of RNA are conserved among all nodaviruses, indicating a selective pressure to maintain the electrostatic interaction with RNA. In BBV these lysines point to weak density, which, given the FHV result, can now be interpreted as RNA also. Only Gln 61 (α I) points to the minor groove of the RNA helix but it is too far away to hydrogen-bond with any bases.

Although the genomes of nodaviruses are single-stranded, the RNA duplex observed here is likely to be part of an extensive secondary structure, which has been observed in other single-stranded RNA viruses¹³. The RNA, together with the peptide arm, stabilizes the flat contact. Neither the peptide arm nor the

FIG. 3 Structure and protein interactions of duplex RNA in FHV. *a*, Stereo view of the native electron density map (thin lines) displayed with the modelled RNA duplex (thick lines). The phases used to calculate this map were derived entirely from modelled protein. The view is down an icosahedral 2-fold axis looking from inside the virus to the exterior. Ten nucleotides were modelled per icosahedral asymmetric unit (NA1-NA10, 5' to 3'). The 2-fold-related strand (NA1'-NA10') forms an RNA duplex. *b*, Stereo view of protein-RNA interactions in FHV. RNA duplex (in bold) is shown with the α -carbon tracing (thin lines) of the first 18 and last 17 residues seen in the crystal structure and the amino-acid side chains (medium lines) that interact with the RNA. All interactions are with the C subunit and its 2-fold-related partner. An icosahedral 2-fold axis is shown with the arrow pointing to the virus centre. Labels with primes indicate 2-fold-related atoms.



RNA are seen at the contact between the A and B₅ subunits, allowing these to form a bent contact. The complex of the RNA and the helix from the cleavage product (γ) may form a more energetically favourable conformation than an uncleaved helix. This would explain how maturation cleavage increases particle stability.

Ordered nucleic acid has so far only been observed in crystal structures of viruses with $T = 1$ capsids^{14–16}. The ordered nucleic acid in all these viruses is single-stranded and although the nucleic acid interacts specifically with the viral capsid, this interaction is not necessary for capsid assembly because all three viruses can form stable empty particles. In contrast, the FHV particle encapsidates RNA as an integral part of the capsid quaternary structure, which explains why empty capsids have not been observed in any nodavirus. □

Received 7 August; accepted 5 November 1992.

- Caspar, D. L. D. *Biophys. J.* **32**, 103–138 (1980).
- Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K. & Bricogne, G. *Nature* **276**, 368–373 (1978).
- Abad-Zapatero, C. et al. *Nature* **286**, 33–39 (1980).
- Hogle, J. M., Maeda, A. & Harrison, S. C. *J. molec. Biol.* **191**, 625–638 (1986).
- Caspar, D. L. D. & Klug, A. *Cold Spring Harbor Symp. quant. Biol.* **27**, 1–24 (1962).
- Hendry, D. A. in *Viruses of Invertebrates* (ed. Kurstak, E.) 227–276 (Dekker, New York, 1991).
- Newman, J. F. E. & Brown, F. J. *gen. Virol.* **21**, 371–384 (1973).
- Mori, K.-I. et al. *Virology* **187**, 368–371 (1992).
- Gallagher, T. M. & Rueckert, R. R. *J. Virol.* **62**, 3399–3406 (1988).
- Kaesberg, P. et al. *J. molec. Biol.* **214**, 423–435 (1990).
- Fernandez-Tomas, C. B. & Baltimore, D. *J. Virol.* **12**, 1122–1130 (1973).
- Rossmann, M. G. & Johnson, J. E. *Rev. Biochem.* **58**, 533–573 (1989).
- Li, T., Chen, Z., Johnson, J. E. & Thomas, G. J. *Biochemistry* **29**, 5018–5026 (1990).
- Chen, Z. et al. *Science* **245**, 154–159 (1989).
- Tsao, J. et al. *Science* **251**, 1456–1464 (1991).
- McKenna, R. et al. *Nature* **355**, 137–143 (1992).
- Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946–950 (1991).
- Smith, T. J. *J. appl. Crystallogr.* **23**, 141–142 (1990).
- Fisher, A. J., McKinney, B. R., Wery, J.-P. & Johnson, J. E. *Acta crystallogr.* **B48**, 515–520 (1992).
- Hosur, M. V. et al. *Proteins: Struct. Funct. Genet.* **2**, 167–176 (1987).
- Arnold, E. et al. *Acta crystallogr.* **A43**, 346–361 (1987).
- Jones, T. A. *J. appl. Crystallogr.* **11**, 268–272 (1978).
- Hogle, J., Kirchhausen, T. & Harrison, S. C. *J. molec. Biol.* **171**, 95–100 (1983).
- Silva, A. M. & Rossmann, M. G. *J. molec. Biol.* **197**, 69–87 (1987).

ACKNOWLEDGEMENTS. We thank B. McKinney for tissue culture work. Ribbon diagrams were generated with the program MOLSCRIPT¹⁷, Fig. 4 with the program MACINPLOT¹⁸, and RNA helical parameters with the program NEWHEL32 (Brookhaven protein data bank; R. E. Dickerson).

Double-helical RNA in satellite tobacco mosaic virus

Steven B. Larson, Stanley Koszelak, John Day,
Aaron Greenwood, J. Allan Dodds
& Alexander McPherson

The Department of Biochemistry and The Department of Plant Pathology,
University of California at Riverside, Riverside, California 92521, USA

SATELLITE tobacco mosaic virus (STMV) is the spherical satellite to an obligatory rod-shaped helper tobacco mosaic virus (TMV), which is required for replication^{1,2}. STMV has 60 protein subunits of M_r 17,500 on a $T = 1$ icosahedral capsid³ containing a single-stranded RNA genome of 1,059 bases^{4–6}. STMV appears similar to another virus, STNV^{7,8}, but is ~20 per cent smaller. It shows no amino-acid homology or immunological cross-reactivity with either STNV or its host TMV^{4,9}. Here we report the X-ray crystal structure of STMV, which shows that the coat protein of STMV contains a 'Swiss roll' β -barrel. An amino-terminal strand extends more than 60 Å and is primarily responsible for quaternary interactions. Each capsid dimer is associated with a segment of genomic RNA double helix comprising seven base pairs. The dyad of each protein dimer is coincident with that of the central base pair of the associated RNA segment whose helix axis is directed along an icosahedral edge. Protein–nucleic acid interactions are extensive. The RNA helices, which have additional stacked bases at their 3' termini, differ significantly from canonical nucleic acid helical forms.

TABLE 1 Structure solution and refinement

Self-rotation function ^{11,12}	Orientation of particle determined by direction of 2-, 3- and 5-fold axes
Rigid-body refinement ³⁶ of STNV atomic model	Resolution: 5–15 Å; result: $R \approx 0.48$; $\langle \Delta \phi \rangle = 83^\circ$; input: atomic model oriented according to self-rotation results
Molecular averaging ^{13,14}	Input: 6–15 Å phases from rigid-body refinement of STNV; result: phases extended to 3 Å; no increase in apparent resolution of maps beyond ~5 Å
Heavy-atom ($F_{\text{der}} - F_{\text{nat}}$) Fourier maps	Input: 5–15 Å phases from molecular averaging; results: two icosahedrally related sets of peaks among 8 derivatives corresponding to two heavy-atom sites about 62 Å from origin
Heavy-atom refinement ³⁷	Input: 15 heavy atoms site per asymmetric unit; 8 derivatives used; resolution: initial = 5–15 Å, final = 3.5–15 Å; result: figure of merit = 0.60 at 5 Å, 0.51 at 3.5 Å; $\langle \Delta \phi \rangle = 92^\circ$ (5–15 Å) from rigid-body STNV model
MIR map	Input: MIR phases for 3.5–15 Å data; result: 55–90 Å spherical band of density around origin; weak elsewhere
Averaged MIR Map ^{13,14}	Input: MIR phases for 3.5–15 Å data; result: 51–90 Å spherical band of strong density traceable for coat protein but with opposite twist of β -sheets from observed twist in other viruses
Molecular averaging phase extension ^{13,16}	Input: phases from heavy atoms with sign of z -coordinate changed; resolution: initial = 3.5–15 Å; final = 2.95–15 Å; envelope: 50–91 Å icosahedrally bounded shell; weights: $w = \exp(-\ F_o\ - \ F_c\ /\ F_o\)$; scaling: F_o not scaled to F_c before each cycle; result: $R = 0.152$; correlation coefficient = 0.953; $\langle \Delta \phi \rangle = 48^\circ$
Averaged electron density map	Result: residues 17–159 identified in minimap by side-chain density; first 16 not apparent but directed towards interior of the particle
FRODO map and model building ²⁴	Result: built capsid protein (residues 17–159) and RNA (6 base pairs + 2 nucleotides); $R = 0.38$; subsequent $2F_o - F_c$ map (averaged) allowed addition of residue 16 and 4 nucleotides to give 7 base pairs + 2 nucleotides
Conjugate gradient minimization ^{18,19}	Parameter sets: protein parameters according to ref. 38 parnahle.dna ^{18,19} for RNA; resolution: 3–6 Å; result: $R = 0.215$; $\langle \Delta \phi \rangle = 54.5^\circ$ (3.5–6 Å) from MIR phases; $\langle \Delta \phi \rangle = 27.1^\circ$ (3–6 Å) from molecular averaged phases

Using multiwire area detector data¹⁰, eight isomorphous derivatives were used to estimate phases to 3.5 Å resolution for the 1222 crystals of STMV ($a = 174.3$ Å, $b = 191.8$ Å and $c = 202.5$ Å). The single orientation ambiguity for the two virus particles in the unit cell was resolved by a self-rotation function¹¹ using MERLOT¹². Phases for 62,474 of 68,034 unique reflections to 3.0 Å were obtained by molecular averaging/phase extension techniques^{13–17}. The structure of the capsid and its associated nucleic acid was refined with maintenance of strict icosahedral symmetry to $R = 0.215$ using XPLOR^{18,19}. A summary of the structure determination is presented in Table 1.

Electron density from 15-fold averaging of the phase-extended MIR map was free of noise, showed high contrast, clear molecular boundaries and excellent connectivity. Antiparallel β -structure was evident and distinctive side chains readily identifiable. The course of the capsid polypeptide chain was traced without interruption from the carboxyl terminus to residue 16. The amino-terminal segment 1–15, containing five arginines and lysines, could not be unambiguously identified and appears to be associated with interior RNA as in most other viruses^{20–22}.

Amino acids 37–159 form a compact eight-stranded antiparallel β -barrel of 'Swiss roll' topology similar to those found in the coat proteins of other viruses^{20–23}. Residues 16–37 comprise a long polypeptide that extends nearly 60 Å from the β -barrel. Interactions between dyad-related subunits, mediated

FIG. 1 Two representative portions of electron density are shown in the icosahedrally averaged, phase-extended MIR Fourier map, which arise from nucleic acid within STMV. *a*, Density belonging to the three central base pairs with structure superimposed; *b*, extended region of this density to which a double-helical RNA segment has been fitted. An icosahedral twofold axis lies in the plane of the central base pair, perpendicular to the helix axis and relating the two antiparallel polynucleotide strands. The helical axis of the segment lies along an edge of the icosahedron.

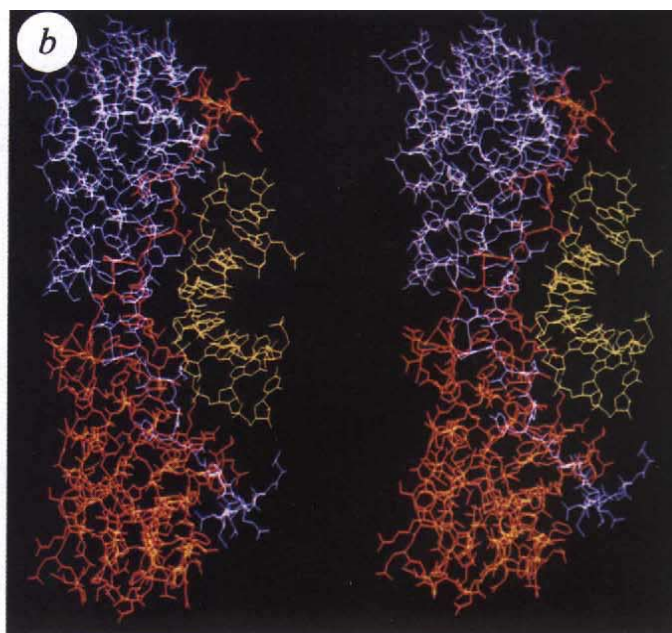
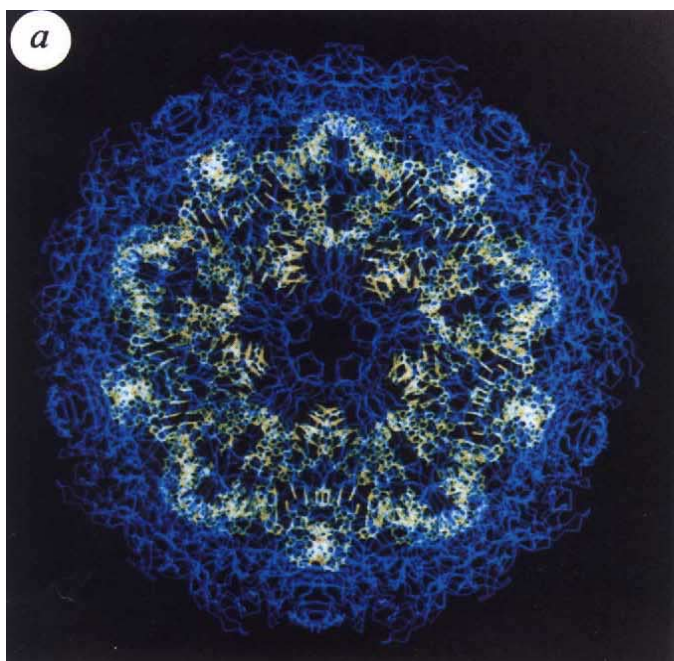
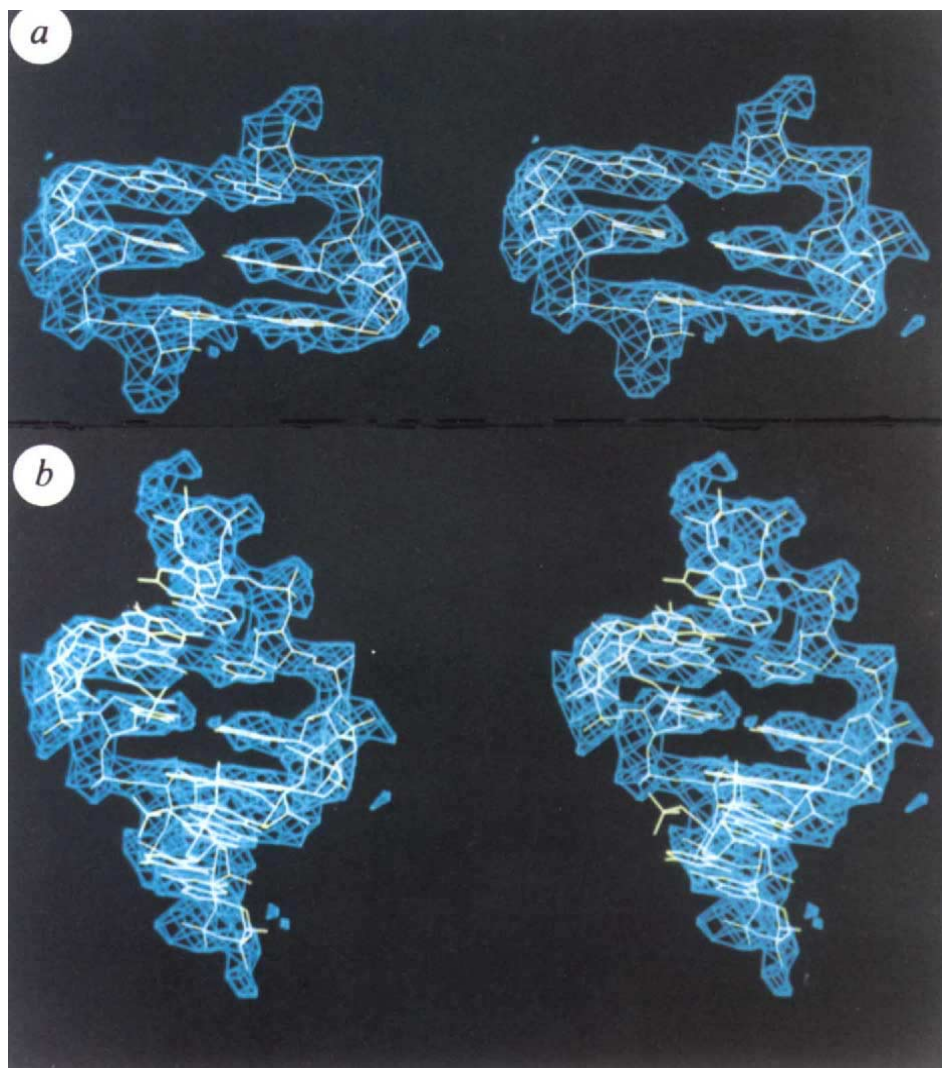
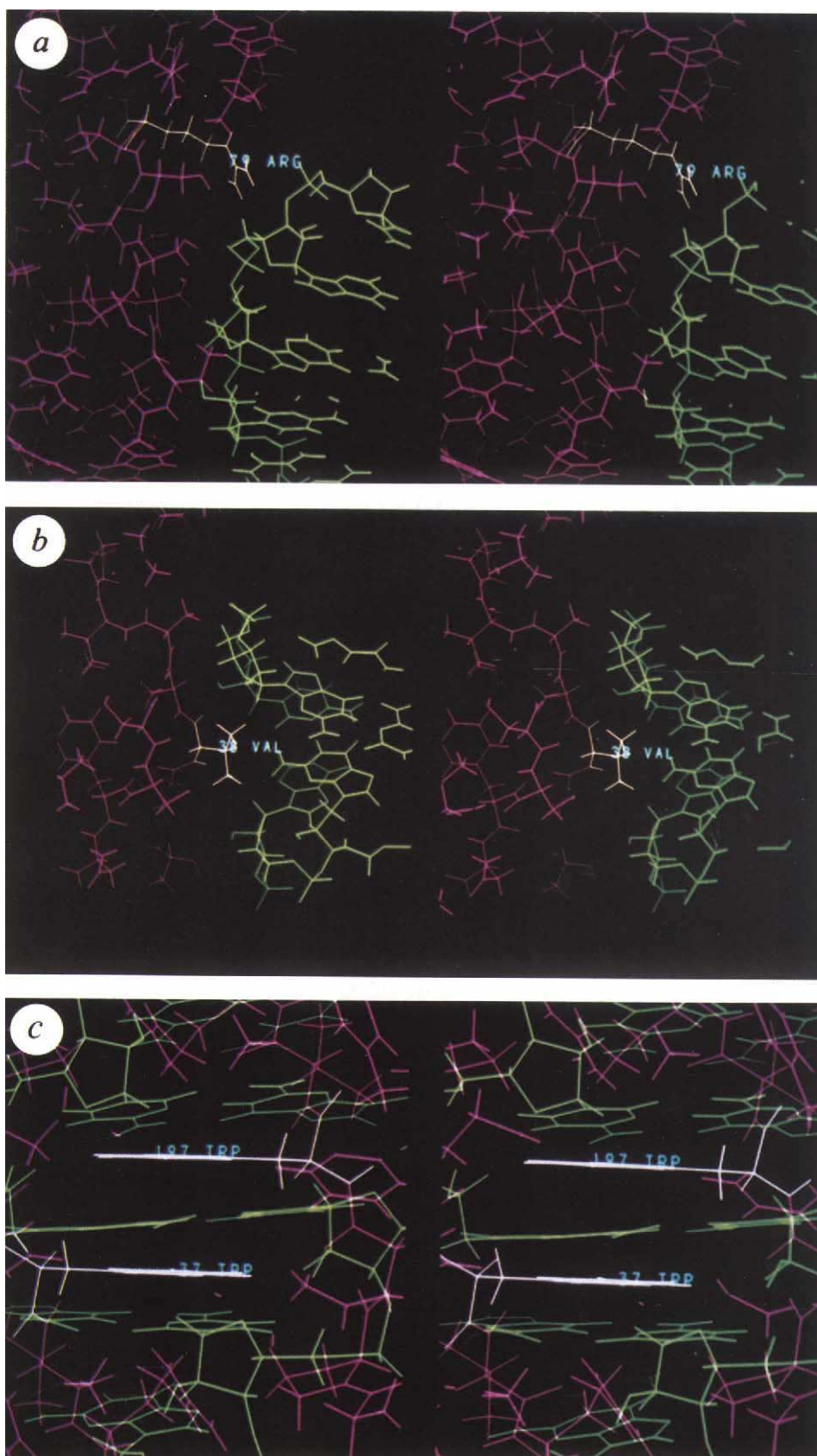


FIG. 3 *a*, Detail of the protein-RNA helix interaction, showing a salt bridge between arginine 79 and nucleic acid phosphates. The protein is represented in purple and the RNA in green. Because of twofold symmetry, every RNA segment forms two such bridges, one with each protein monomer, as for all interactions involving the helices. *b*, A valine side chain at position 38 is prominently placed in the minor groove of the helix. The most intriguing constellation of side chains at the protein-RNA interface in *c* involves twofold-related tryptophan-37 side chains (197 designates tryptophan 37 in the associated subunit). These are stacked upon one another with 3.6 Å separation. The curious aspect of this formation is that the planes of the tryptophan rings are parallel with the plane of the central base pair of the associated RNA helix.



◀ FIG. 2 *a*, The STMV protein capsid (in blue) is associated with thirty icosahedrally distributed RNA double helical segments (yellow). Each segment contains 16 nucleotides and 7 base pairs. The protein forms a shell between radii 57 Å and 86 Å, whereas those RNA segments visible in the electron density map, comprising about 40–45% of the STMV genome, lie between 52 Å and 64 Å radius. The diameter of STMV is less than STNV, with radii at threefold axes of 80 Å and 86 Å respectively, and 86 Å versus 96 Å at the fivefolds where the prominent cusps of STNV are almost eliminated in STMV. *b*, All non-hydrogen atoms of two subunits of the coat protein of STMV, these are related by twofold symmetry (one copy in red and the other in purple). The associated double-helical RNA segment of sixteen nucleotides is shown in yellow. The view is perpendicular to a horizontal icosahedral dyad axis which is exact for both the protein and the nucleic acid. Note the extended amino-terminal strands that bind protein subunits together while forming protein-RNA interactions as well.

principally by this amino-terminal strand, are so extensive that it is almost certain that dimers of protein must assemble to form the capsid. In addition the extended terminal strand provides most of the threefold interactions.

Once capsid polypeptides were accounted for, a substantial portion of well defined electron density remained inside the virion and positioned on icosahedral dyads. The density took the form of paired antiparallel arcs, crossbridged by planar bands of density, with each arc appearing as a sequential array of strong peaks at constant distances of 5.8 Å but connected by weaker density. The pattern was unmistakably that of base-paired double-helical nucleic acid. Because the map had been averaged, the distribution of the double-helical segments was consistent with icosahedral symmetry. The helical axes of the RNA segments lay exactly in the planes defined by neighbouring fivefold axes, that is, parallel to edges of the icosahedron.

An RNA model using G-C base pairs and 3'-endo ribose was built using computer graphics²⁴. The electron density, illustrated in Fig. 1, accommodated two antiparallel strands, each composed of eight nucleotides. Seven base pairs were clearly evident, with one apparently unpaired, but stacked nucleotide at the 3' end of each strand in the helix. The helices lay precisely on viral twofold axes of symmetry, so that twofold axes of the RNA segments and icosahedral dyads were coincident. The twofold axes lay exactly in the planes of the central base pairs, perpendicular to the RNA helical axes, thereby relating base-paired polynucleotide strands.

The STMV genome was analysed for its capacity to self-base pair using the program FOLD^{25,26}. The single-stranded RNA of STMV has an exceptional potential for forming Watson-Crick base pairs and secondary structure compatible with double-helical RNA. From the 68% of STMV bases predicted to be paired, adequate candidate segments could readily be found to fill the thirty binding sites necessary.

A segment of RNA from STMV can be extended to create a model helix of arbitrary length. Parameters for such a helix were obtained using the program NEWHELIX²⁷. The base tilt is about 70°, similar to A-form DNA, and that observed in transfer RNA structures^{28,29}. The RNA helix is otherwise more similar to B-form DNA, with a pitch of 30–32 Å, 10 base pairs per turn, and a diameter of about 17 Å. Width and depth of the major and minor grooves more closely resemble B-form DNA. The distribution of the 30 RNA segments on the interior of the virus is shown in Fig. 2a.

Double helical RNA segments are closely associated with the interior surface of the capsid, each in intimate contact with a protein dimer with which it shares a common twofold axis. The capsid dimer thus constitutes a helical-RNA-binding protein. Two protein subunits, seen in Fig. 2b, create a cradle at their juncture on the inside surface of the capsid. RNA segments lie crosswise in the cradles. The van der Waals surfaces of protein and nucleic acid are in direct contact, although the interface could incorporate water molecules.

The observed RNA-protein contacts arise primarily from main- and side-chain atoms of several strands of the β -barrels and from residues 16–37 in the amino-terminal polypeptide. The fifteen amino-terminal residues must, we believe, penetrate further into the interior where they associate with other, less ordered RNA. We have not yet delineated all of the important protein-nucleic acid interactions, but some are illustrated in Fig. 3. We currently see at least four twofold-related pairs of salt bridges between arginine and lysine side chains and RNA phosphates. Many other opportunities for hydrogen bonds exist as well.

Nucleic acid has been previously observed in icosahedral virus structures by X-ray crystallography^{30–32}, and ordered nucleic acid has been indicated by other physical methods as well^{33,34}. Stacked bases of single-stranded DNA were observed in canine parvovirus³⁰ and some partially ordered DNA single strands in Φ X174 (ref. 31). In each case they accounted for about 12 per cent of the total DNA bases. Stacked nucleotides

of single-stranded RNA were also observed in bean pod mottle virus³². The double-stranded RNA seen in STMV, however, represented the highest proportion of the genome, 44%, of any virus so far visualized. It further imposes a stringent constraint on the organization of the genome and its distribution within the virus. Diffraction data to 1.8 Å resolution are in hand for the STMV crystals³⁵ and high-resolution refinement may provide additional details regarding connectivity of the helical RNA segments. □

Received 29 June; accepted 13 November 1992.

- Valverde, R. A. & Dodds, J. A. *J. gen. Virol.* **67**, 1875–1884 (1986).
- Valverde, R. A. & Dodds, J. A. *J. gen. Virol.* **68**, 965–972 (1987).
- Casper, D. L. & Klug, A. *Cold Spring Harbor Symp. quant. Biol.* **27**, 1–24 (1962).
- Mirkov, T. E., Mathews, D. M., DuPlessis, D. H. & Dodds, J. A. *Virology* **170**, 139–146 (1989).
- Kim, K. S., Valverde, R. A. & Dodds, J. A. *J. ultrastr. Molec. Struct. Res.* **102**, 196–204 (1989).
- Kozzelak, S., Dodds, J. A. & McPherson, A. *J. molec. Biol.* **209**, 323–325 (1989).
- Liljas, L. *et al. J. molec. Biol.* **159**, 93–108 (1982).
- Jones, A. T. & Liljas, L. *J. molec. Biol.* **177**, 735–767 (1984).
- Dodds, J. A. *Can. J. Plant Path.* **13**, 192–195 (1991).
- Xuong, N.-H., Nielson, C., Hamlin, R. & Anderson, D. *J. appl. Crystallogr.* **18**, 342–360 (1985).
- Rossmann, M. G. & Blow, D. M. *Acta crystallogr.* **15**, 24–31 (1962).
- Fitzgerald, P. M. D. *J. appl. Crystallogr.* **21**, 273–276 (1988).
- Bricogne, G. *Acta crystallogr.* **A30**, 395–405 (1974).
- Bricogne, G. *Acta crystallogr.* **A32**, 832–847 (1976).
- Arnold, E. *et al. Acta crystallogr.* **A43**, 346–361 (1987).
- Arnold, E. & Rossmann, M. G. *Acta crystallogr.* **A44**, 270–282 (1988).
- Rayment, I. *Acta crystallogr.* **A39**, 102–116 (1983).
- Brunger, A. T. *Rev. phys. Chem.* **42**, 197–223 (1991).
- Brunger, A. T., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).
- Rossmann, M. G. & Johnson, J. E. *Rev. Biochem.* **58**, 533–573 (1989).
- Branden, C. & Tooze, J. *Introduction to Protein Structure* Ch. 11 (Garland, New York 1991).
- Abad-Zapatero, C. *et al. Acta crystallogr.* **B37**, 2002–2018 (1981).
- Branden, C. & Tooze, J. *Introduction to Protein Structure* 70–72, (Garland, New York, 1991).
- Jones, T. A. in *Computational Crystallography* (Sayre, D. ed.) 307–317 (Oxford Univ. Press, New York, 1982).
- Devereux, J., Haeblerli, P. & Smithies, O. *Nucleic Acids Res.* **12**, 387–395 (1984).
- Zucker, M. & Stiegler, P. *Nucleic Acids Res.* **9**, 133–148 (1981).
- Dickerson, R. E., Kopka, M. L. & Pjura, P. in *Biological Macromolecules and Assemblies, Vol. II, Nucleic Acids and Interactive Proteins* (eds Jurnak, F. A. & McPherson, A.) (Wiley, New York, 1985).
- Kim, S. H. *et al. Science* **185**, 435–440 (1974).
- Robertus, J. D. *et al. Nature* **250**, 546–551 (1974).
- Tsao, J. *et al. Science* **251**, 1456–1464 (1991).
- McKenna, R. *et al. Nature* **355**, 137–143 (1992).
- Chen, Z. *et al. Nature* **245**, 154–159 (1989).
- Bantley, G. A. *et al. J. molec. Biol.* **194**, 129–141 (1987).
- Benevides, J. M., Stow, P. L., Leodevico, L. I., Incardona, N. L. & Thomas, G. J. *Biochemistry* **30**, 4855–4863 (1991).
- Day, J. & McPherson, A. *Protein Sci.* **1**, 1254–1268 (1992).
- Tronrud, D. E., Ten Eyck, L. F. & Mathews, B. W. *Acta crystallogr.* **A43**, 489–501 (1987).
- Terwilliger, T. & Eisenberg, D. *Acta crystallogr.* **A39**, 813–816 (1984).
- Engh, T. A. & Huber, R. *Acta crystallogr.* **A47**, 392–400 (1991).

ACKNOWLEDGEMENTS. This research was supported by grants from NIH, NSF and NASA. We thank J. Heick and M. Angel for preparing the virus, T.-P. Ko, N. Ban and M. Greene for assistance and advice, and the San Diego Supercomputer Center for time on the Cray Y-MP.

Evolution *in vitro* of an RNA enzyme with altered metal dependence

Niles Lehman & Gerald F. Joyce

Departments of Chemistry and Molecular Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037, USA

THE *Tetrahymena* group I ribozyme catalyses a sequence-specific phosphodiester cleavage reaction on an external RNA oligonucleotide substrate in the presence of a divalent metal cation cofactor¹. This reaction proceeds readily with either Mg²⁺ or Mn²⁺, but no detectable reaction has been reported when other divalent cations are used as the sole cofactor^{2–5}. Cations such as Ca²⁺, Sr²⁺ and Ba²⁺ can stabilize the correct folded conformation of the ribozyme, thereby partially alleviating the Mg²⁺ or Mn²⁺ requirement². But catalysis by the ribozyme involves coordination of either Mg²⁺ or Mn²⁺ at the active site, resulting in an overall requirement for one of these two cations⁵. Here we use an *in vitro* evolution process^{6,7} to obtain variants of the *Tetrahymena* ribozyme that are capable of cleaving an RNA substrate in reaction mixtures

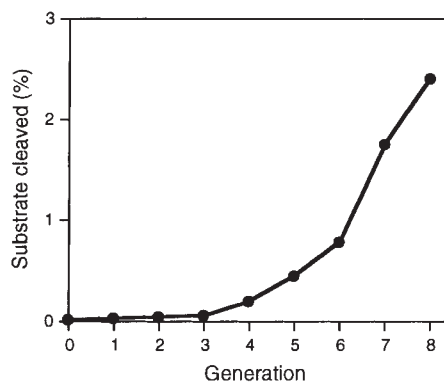


FIG. 1 Improvement over successive generations in the ability of the evolving population of ribozymes to carry out sequence-specific phosphodiester cleavage of an external RNA substrate in 10 mM CaCl_2 .

METHODS. The starting population was generated using four mutagenic oligodeoxynucleotides that randomize 140 nucleotide positions encompassing the catalytic centre of the ribozyme at an expected degeneracy of 5% mutations per position⁷. The selection protocol was as described⁷, but with the following modifications. Conditions for RNA-catalysed RNA cleavage:

40 pmol ribozyme (2.4×10^{13} molecules), 100 pmol substrate (5'-CCCUCU-A₃UA₃UA-3'), 10 mM CaCl_2 and 30 mM EPPS, pH 7.5 in 40 μl were incubated at 37 °C for 3.25 h. Ribozyme products were purified by precipitation with ethanol, electrophoresis through a 20% polyacrylamide/8 M urea gel (to preclude reaction with substrate during subsequent manipulations in Mg^{2+} -containing buffers), and affinity chromatography on Nensorb 20 (New England Nuclear). The entire population of purified ribozymes was amplified by selective isothermal RNA amplification^{20,21} in a reaction mixture containing 10 mM MgCl_2 , 80 mM CH_3COOK , 5 mM dithiothreitol, 50 mM Tris-HCl, pH 7.5, 2 mM NTPs, 0.2 mM dNTPs, 50 pmol DNA primers, 250 U MMLV reverse transcriptase and 200 U T7 RNA polymerase in 50 μl at 37 °C for 2 h. Subsequent selective cDNA synthesis used 10% of the selective amplification product in 10 mM MgCl_2 , 5 mM dithiothreitol, 50 mM Tris-HCl, pH 7.5, 2 mM NTPs, 0.2 mM dNTPs, 20 pmol DNA primer and 20 U AMV reverse transcriptase in 20 μl at 42 °C for 30 min. Subsequent non-selective PCR used 5% of the selective cDNA synthesis product. PCR-generated double-stranded DNA was transcribed in the presence of [³H]UTP. RNA products were purified by electrophoresis through a 5% polyacrylamide/8 M urea gel and by Nensorb chromatography. Conditions for assay of RNA-catalysed RNA cleavage: 1 μM ribozyme, 2.5 μM substrate (5'-GGCCUCU-A₃UA₃UA-3'), prepared by *in vitro* transcription²² in the presence of [α -³²P]ATP, 10 mM CaCl_2 , 30 mM EPPS, pH 7.5 in 7.5 μl at 37 °C for 3.25 h. Reaction products were separated by electrophoresis through a 20% polyacrylamide/8 M urea gel and then quantified by Cerenkov counting.

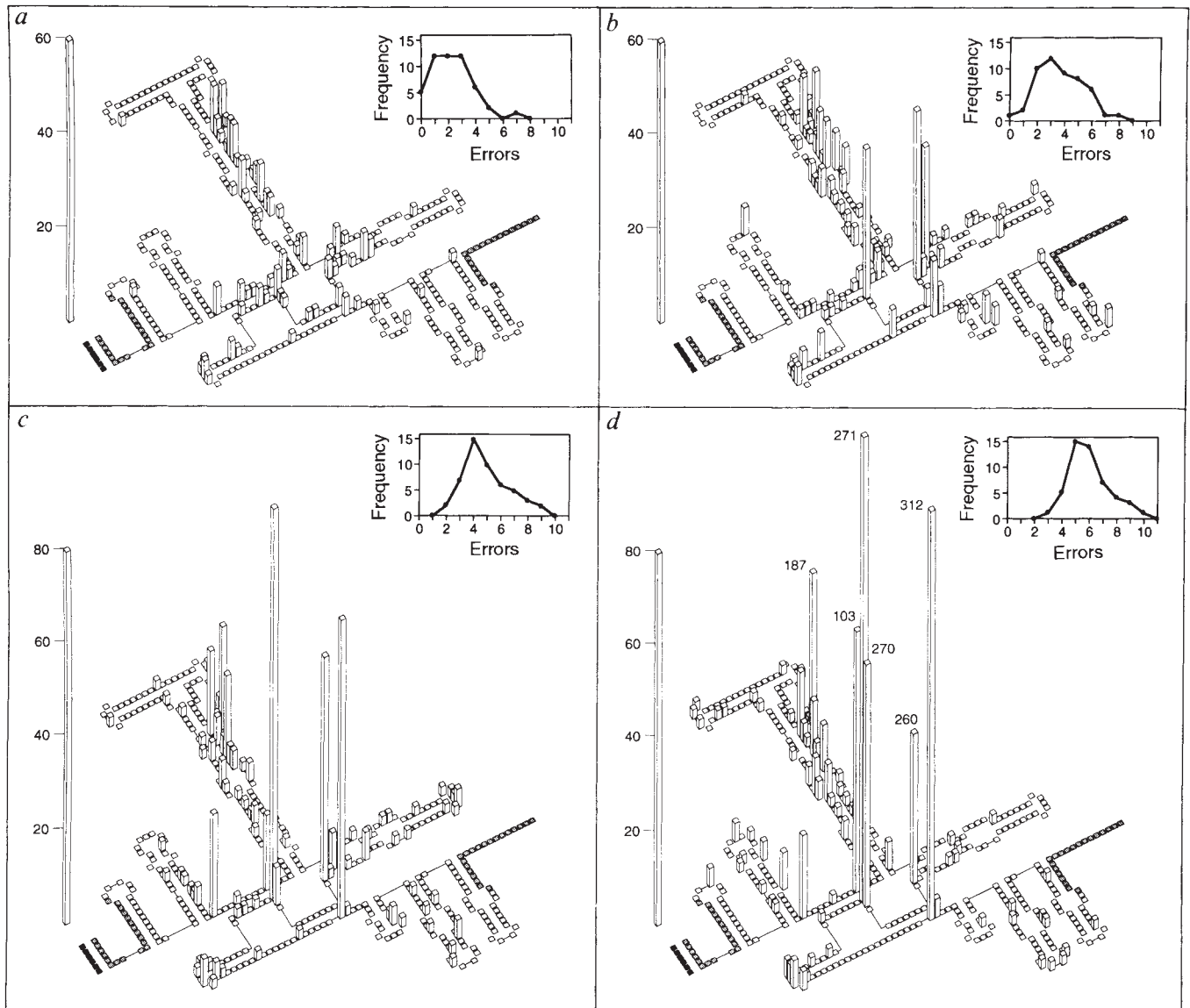


FIG. 2 Sites at which mutations occurred over the course of evolution, superimposed on the secondary structure of the *Tetrahymena* ribozyme. Box height corresponds to frequency of mutation (per cent, based on 50 clones). Immutable primer binding sites are shaded; substrate is shown in black. Insets detail the number of clones with a particular number of

mutations relative to the wild-type sequence. *a*, Generation 2; *b*, generation 4; *c*, generation 6; *d*, generation 8. Labelled positions in *d* are sites of the most frequent mutations. Nucleotide sequences of individual clones are available on request.

containing Ca^{2+} as the divalent cation. These findings extend the range of different chemical environments available to RNA enzymes and illustrate the power of *in vitro* evolution in generating macromolecular catalysts with desired properties.

We subjected a starting pool of 2.4×10^{13} *Tetrahymena* ribozyme variants with partially randomized sequences to eight cycles of selection, amplification and mutation (eight generations). In each generation those individual variants that catalysed cleavage of a target RNA substrate under conditions of 10 mM CaCl_2 were preferentially amplified. The successive populations were assayed for their composite ability to cleave a radioactively labelled substrate in 10 mM CaCl_2 (Fig. 1). The trace amount of activity that we detect for the wild-type *Tetrahymena* ribozyme is significantly above background only after a 3-hour incubation and, accordingly, we chose a 3.25-hour incubation time for our selection protocol. Our starting population (G_0) does not display any activity, but by the eighth generation (G_8) the evolving population is 170-fold more active than the wild type (Fig. 1).

To determine which mutations were responsible for the continuously improving phenotype, we obtained 50 individual clones from each of generations 2, 4, 6 and 8 for direct sequence analysis (Fig. 2). Mutations were found to occur throughout the 353 variable nucleotide positions of the ribozyme, with most, but not all, falling within the central 140-nucleotide region that was most extensively randomized in the starting population. Many positions in this central region did not have any mutations in the 200 clones that were sequenced. Presumably these sites require specific nucleotide bases for the ribozyme to be catalytically active, a result consistent with comparative phylogenetic⁸⁻¹² and site-directed mutational¹³⁻¹⁷ analyses. By contrast, six nucleotide positions are mutated in more than 30% of the G_8 clones (Fig. 2). Specifically, a U $\rightarrow$ C mutation at position 271 is fixed by generation 8, occurring in all 50 clones that were sequenced. The other frequent mutations in G_8 are a G $\rightarrow$ A change at position 312 (88%), an A $\rightarrow$ G change at position 103 (56%), an A $\rightarrow$ G change at position 270 (52%), an A $\rightarrow$ U, C, G change at position 187 (34%), and a C $\rightarrow$ A change at position 260 (31%). Of these mutations, all but C $\rightarrow$ A at position 260 increase steadily in frequency over the eight generations. Only one, G $\rightarrow$ A at position 312, was detected in more than 10% of 50 clones from the ninth generation of a previous study aimed at selecting ribozymes competent at cleaving a DNA substrate⁷.

The high frequencies of these six mutations suggested that they are selectively advantageous. When the co-occurrences of two or more specific mutations were considered, it became clear

that the evolving population was converging on two distinct classes of solutions, one containing the C $\rightarrow$ A mutation at position 260, and the other containing the A $\rightarrow$ G mutation at position 270. These two mutations do not occur together in any of the clones that were sequenced.

To probe the functional consequences of particular mutations and combinations of mutations, we prepared 21 individual RNAs from sequenced clones and tested their catalytic properties (Figs 3 and 4). These clones include four single-error mutants (at positions 189, 260, 271 and 312), and 16 higher-error mutants containing various combinations of the six high-frequency mutations, with or without auxiliary mutations. When assayed in 10 mM CaCl_2 , the higher-error mutants are consistently more reactive than the lower-error mutants. The least reactive individuals are the four single-error mutants.

Of all the individual variants that we studied, the two most reactive are high-error mutants, each representing one of the two mutually exclusive classes of solutions: clone 41 (from G_6), having mutations at positions 94, 187, 225, 260, 271 and 312, and clone 61 (from G_8), having mutations at positions 103, 187, 270, 271 and 312. Clone 61 is one of five G_8 clones that have this exact sequence. No other specific sequence, other than the wild type in G_2 , is ever observed more than once. In fact, the 103/270/271/312 combination is present in 42% of G_8 clones.

On the basis of the fraction of substrate cleaved after 3.25 hours in 10 mM CaCl_2 , clones 41 and 61 are, respectively, 243-fold and 301-fold more potent than the wild type. Although the activity of the wild type in 10 mM CaCl_2 is too low to permit formal kinetic analysis, catalytic efficiency, k_{cat}/K_m , could be determined for clones 41 and 61 as $10,400$ and $8,550 \text{ M}^{-1} \text{ min}^{-1}$, respectively (see legend to Fig. 4). Catalytic efficiency of the wild type in a related reaction in 10 mM MgCl_2 at 50°C was reported as $9 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$ (ref. 18).

There is a weak but significant correlation between the activity of an individual in 10 mM CaCl_2 and its activity in 10 mM MgCl_2 ($r = 0.61$, $P < 0.05$). All of the evolved individuals that we studied rank higher than the wild type in both the Ca^{2+} and Mg^{2+} comparisons (Figs 3 and 4).

The ribozymes that we have discovered as a result of selection for activity in Ca^{2+} alone remain efficient catalysts in Mg^{2+} (Fig. 4) and Mn^{2+} (data not shown). Thus, three possibilities must be considered: (1) that they have become modified in such a way as to accommodate Ca^{2+} in place of Mg^{2+} or Mn^{2+} at the active site; (2) that they are able to scavenge minute amounts of Mg^{2+} and/or Mn^{2+} that exist as impurities in our reaction mixtures, but cannot use Ca^{2+} at the catalytic site; (3) that they

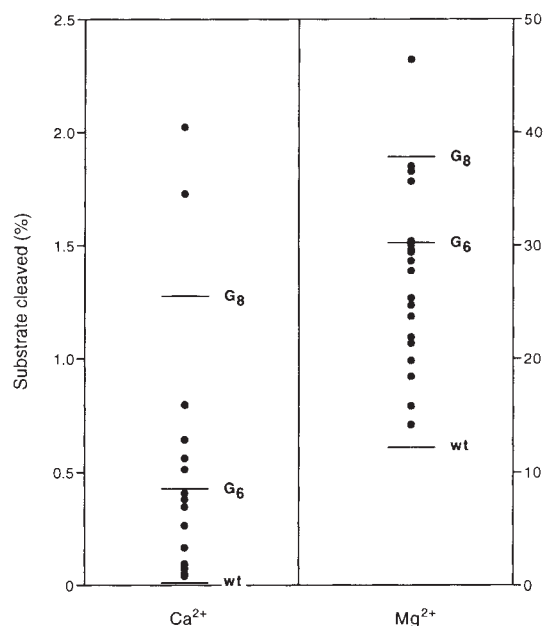
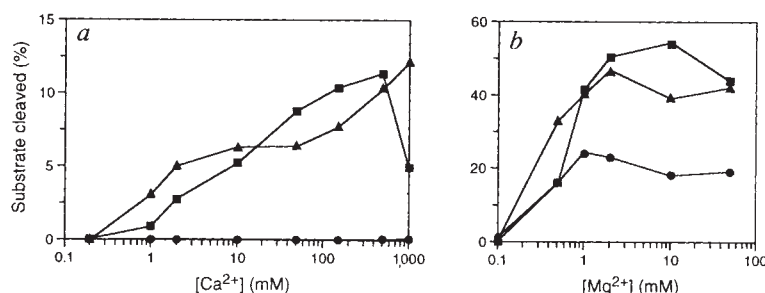


FIG. 3 Catalytic activity of 20 individual ribozyme variants obtained through generation 6, assayed in a, CaCl_2 , or b, MgCl_2 (clone 61 from G_8 not included). Horizontal lines labelled G_6 and G_8 refer to the activity of the entire population at generations 6 and 8, respectively. wt, Wild-type activity. Reaction conditions: $0.27 \mu\text{M}$ ribozyme, $0.67 \mu\text{M}$ [$\alpha\text{-}^{32}\text{P}$] ATP-labelled substrate, 10 mM CaCl_2 or MgCl_2 , 30 mM EPPS, pH 7.5 in $7.5 \mu\text{l}$ at 37°C for 3.25 h.

FIG. 4 Catalytic activity of the wild type and two ribozyme variants in the presence of varying concentrations of a, CaCl_2 or b, MgCl_2 . Reaction conditions: 0.27 μM ribozyme, 0.67 μM [$\alpha\text{-}^{32}\text{P}$] ATP-labelled substrate, 30 mM EPPS, pH 7.5, in 7.5 μl at 37 °C for 3.25 h. ●, Wild type; ■, clone 41; ▲, clone 61. Both clones were inactive in either 10 mM SrCl_2 or 10 mM $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2$ (data not shown). CaCl_2 solution was prepared from >99.99% calcium chloride hydrate solid (Aldrich) and deionized water purified through a Milli-Q UF-plus water ultrafiltration apparatus (Millipore). Metal ion contamination was assayed by inductively coupled plasma emission spectrometry, which detected 0.4 μM Mg and 0.01 μM Mn following incubation at 37 °C for 3.25 h. Under conditions of ribozyme excess (see below), catalytic activity did not increase proportionally with increasing ribozyme concentration above 50 nM (data not shown), indicating inconsequential metal ion contamination of the ribozyme preparations. Catalytic efficiency, k_{cat}/K_m , was determined with (5'- ^{32}P)-labelled substrate in the presence of excess ribozyme: 1 nM substrate, 10 nM or



25 nM ribozyme, 10 mM CaCl_2 , 30 mM EPPS, pH 7.5, at 37 °C for 20 min, 2 h, 4 h and 7 h. Initial rate was linear over this time period ($r_{\text{min}}=0.97$, $r_{\text{avg}}=0.99$). A plot of rate versus ribozyme concentration was linear between 0 and 25 nM ribozyme ($r_{\text{avg}} > 0.99$), the slope providing an estimate of k_{cat}/K_m .

no longer require any divalent cation at the catalytic site, but remain dependent on divalent metals for folding. The second alternative is supported by our findings that the wild type has a trace level of activity in CaCl_2 and that all of the evolved individuals that we studied are improved in MgCl_2 compared to the wild type. But the combined Mg^{2+} and Mn^{2+} impurity in our reaction buffer was measured at 0.4 μM , whereas doping of the 10 mM CaCl_2 reaction mixture with increasing amounts of MgCl_2 did not alter the catalytic activity of the clones below 10 μM added MgCl_2 (see legend to Fig. 4). Furthermore, we detected a behavioural shift that indicated a dependence on Ca^{2+} of the evolved clones: the wild type is more reactive in a buffer containing 0.1 mM MgCl_2 than in a buffer containing 10 mM CaCl_2 and 0.1 mM MgCl_2 (2.5% versus 1.5% substrate cleaved), but clone 61 is less reactive in 0.1 mM MgCl_2 alone than in 10 mM CaCl_2 plus 0.1 mM MgCl_2 (8.9% versus 29.5%). We cannot exclude the third alternative, that the ribozymes have become metal-independent at the catalytic site.

The artificial phylogeny that we have described represents a dramatic example of engendering an informational macromolecule with a new phenotype by *in vitro* evolution. A temporal trace of the evolving sequences bears similarity to patterns of gene-frequency changes in natural populations under strong directional selection. After only eight generations of selective amplification, a highly diverse mutant pool converged onto two similar but mutually exclusive solutions. Further propagation of the phylogeny will determine whether one of these solutions drives the other to extinction and whether clones superior to clones 41 and 61 can be found. It is notable that our populations did not behave according to Fisher's fundamental theorem of natural selection¹⁹, which states that "the rate of increase in fitness of any population at any time is equal to its genetic variance in fitness at that time." Those populations with the greatest variance, the earlier generations, showed less rapid phenotypic improvement than the later, less variant populations (Fig. 1). This situation is probably a reflection of the genotypic plasticity of the *Tetrahymena* ribozyme; much of the variation that was introduced at the outset of this study may have had little effect on the ribozymes' capacity to meet the imposed selection constraint.

- Michel, F. & Dujon, B. *EMBO J.* **2**, 33–38 (1983).
- Cech, T. R. *Gene* **73**, 259–271 (1988).
- Michel, F. & Westhof, E. *J. molec. Biol.* **216**, 585–610 (1990).
- Price, J. V., Kieft, G. L., Kent, J. R., Sievers, E. L. & Cech, T. R. *Nucleic Acids Res.* **13**, 1871–1889 (1985).
- Szostak, J. W. *Nature* **322**, 83–86 (1986).
- Burke, J. M. *Gene* **73**, 273–294 (1988).
- Couture, S. et al. *J. molec. Biol.* **215**, 345–358 (1990).
- Beaudry, A. A. & Joyce, G. F. *Biochemistry* **29**, 6534–6539 (1990).
- Herschlag, D. & Cech, T. R. *Biochemistry* **29**, 10159–10171 (1990).
- Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930).
- Guatelli, J. C. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 1874–1878 (1990).
- Joyce, G. F. in *Antisense RNA and DNA* (ed. Murray, J. A. H.) 353–372 (Wiley-Liss, New York, 1992).
- Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. *Nucleic Acids Res.* **15**, 8783–8798 (1987).

ACKNOWLEDGEMENTS. We thank D. Decker for discussion and technical support, W. Paplawski and G. Arrhenius for plasma emission spectrometry, and D. Herschlag for advice on kinetic analysis. This work was supported by grants from NASA and the NIH.

GUIDE TO AUTHORS

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without interruptions, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. High priority cannot be given to pre-submission enquiries, in urgent cases they can be made in the form of a one-page fax. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers, limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are typeset from the London office.

Nature requests authors deposit sequence and crystallographic data in the databases that exist for this purpose, and to mention availability of these data.

Once a manuscript is accepted for publication, contributors will receive proofs in about 4 weeks. *Nature's* staff will edit manuscripts with a view to brevity and clarity, so contributors should check proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review articles survey recent developments in a field. Most are commissioned but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarize its results and implications. Articles have fewer than 50 references and may contain a few short subheadings.

Letters are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and should contain fewer than 10 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned but suggestions can be made to the Commentary Editor in the form of a one-page synopsis.

News and Views articles inform non-specialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the News and Views Editor.

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and four photocopies are required, each accompanied by artwork. If photographs are included, five sets of originals are required; for line drawings, one set of originals and four good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Three copies of relevant manuscripts in press or submitted for publication elsewhere should be included with each copy of a submitted manuscript, and clearly marked as such. Five copies of revised and resubmitted manuscripts, labelled with their manuscript numbers are required, together with five copies of a letter detailing the changes made.

Titles are brief and simple. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Abstract should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Lettered originals of photographs should be provided. Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome. Original artwork is returned when a manuscript cannot be published.

Protein/nucleic acid sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 100 amino acids or 60 base pairs.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference should be cited in the text as a personal communication, manuscript submitted or in preparation.

Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963–65). The first and last page numbers are included; reference to books should include publisher, place and date.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used in the text.

Acknowledgements are brief and appear after the reference list; grant and contribution numbers are not allowed.

Supplementary information is material relevant to Articles or Letters which cannot, for lack of space, be published in full, but which is available from *Nature* on request.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1234 National Press Building, Washington, DC 20045, USA. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax.

Received 21 September; accepted 20 November 1992.

- Zaug, A. J. & Cech, T. R. *Science* **231**, 470–475 (1986).
- Grosshans, C. A. & Cech, T. R. *Biochemistry* **28**, 6888–6894 (1989).
- Pyle, A. M., McSwiggen, J. A. & Cech, T. R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 8187–8191 (1990).
- Celander, D. W. & Cech, T. R. *Science* **251**, 401–407 (1991).
- Piciorilli, J. A., Vyle, J. S., Caruthers, M. H. & Cech, T. R. *Nature* **361**, 85–88 (1993).
- Joyce, G. F. *Gene* **82**, 83–87 (1989).
- Beaudry, A. A. & Joyce, G. F. *Science* **257**, 635–641 (1992).
- Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A. & Scazzocchio, C. *Nature* **300**, 719–724 (1982).
- Michel, F., Jacquier, A. & Dujon, B. *Biochimie* **64**, 867–881 (1982).

Biospecific interaction analysis using biosensor technology

Magnus Malmqvist

Fundamental information that enhances our understanding of biospecific interactions can be obtained using a new analytical system based on biosensor technology. The functional characteristics of biospecific interaction, such as kinetics, affinity and binding position, are examined by label-free analysis of proteins in free solution binding to an immobilized ligand at a hydrophilic sensor surface.

INTERACTIONS between biomolecules are fundamental for life and the characterization of these interactions in terms of structure-function relationships is crucial for our understanding of biological systems. A new biosensor-based analytical system, BIAcore (Pharmacia Biosensor AB, Uppsala, Sweden), is designed for functional characterization of protein-protein, protein-DNA and ligand-receptor interactions in real time. To date, the analytical system has mostly been used for the kinetic analysis of monoclonal and recombinant antibody-antigen interactions and for biomolecular engineering and drug design.

SPR technology

The adsorption of biomolecules to an immobilized ligand on a sensor chip is measured in the same time and place as it occurs. The analytical system, BIAcore, is based on a biosensor that uses surface plasmon resonance (SPR)¹ to monitor the adsorption of biomolecules on a sensor chip. This optical technique measures changes in refractive index in the vicinity of the surface. Such changes are directly proportional to the change in adsorbed mass², which makes it suitable for the detection of biomolecules. The system includes a sensor chip to which

the ligand can be immobilized in a 100-nm thick hydrophilic matrix composed of 2–3 per cent flexible dextran^{3,4}, a miniaturized fluidics cartridge for the transport of analytes and reagents to the sensor chip⁵, a SPR detector, an auto injector and software for system control and evaluation of results. The system has been described in more detail in refs 6 and 7 (see Fig.1).

Specific ligands are immobilized to the sensor chip at low concentration through biotin-avidin interaction or covalently through amine, thiol or aldehyde chemistry. This provides the user with flexibility in both selecting specificity and in the analytical procedure. The stable binding allows regeneration of the sensor surface and 50–100 analytical cycles can be performed on one and the same surface — depending on regeneration conditions. An easy to use programming environment for automating analytical procedures allows the system to run overnight and at weekends, leaving the daytime free for developing new analyses and evaluation of results. The system can also be used for standardized concentration analysis^{8–10}, as well as for more sophisticated analysis of kinetics and affinity.

The analysis can also be performed in tissue culture media¹¹ or bacterial broth^{12,13} without the need for purification, in contrast to other techniques used for kinetics analysis such as stopped-flow fluorescence quenching.

Applications

Kinetics. The adsorption of an analyte to a surface-immobilized ligand under constant flow can be regarded as a 'pseudo' first-

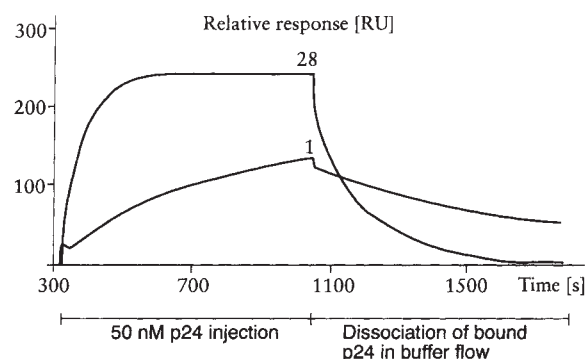


FIG. 2 Sensorgrams for adsorption of HIV-1 p24 antigen on two different mAbs. The curves illustrate the differences in association and dissociation rates and the time difference to reach equilibrium.

order reaction as the concentration of analyte in the thin layer cell is constant. The adsorption rate can be measured using BIAcore^{14–16}, and from these data the association rate constants can be obtained by carrying out a series of adsorption experiments with different analyte concentrations. The logarithmic decrease in complex concentration on the surface after the sample pulse has passed can be used for the determination of dissociation rate constants.

The values obtained from kinetics and affinity measurements based on different principles and methods may differ. However, with an understanding of the fundamental differences of methods, results generated can provide useful quantitative information of interaction parameters. Using BIAcore, the analysis of nonpurified samples can be performed without labels.

Kinetic analysis of mAb binding. Kinetic information about biospecific interactions can complement affinity measurements. For example, two monoclonal antibodies (mAbs) with the same affinity for the antigen had a five-fold difference in association and dissociation rate constants¹⁴ (see Fig. 2). Similar results were obtained for site-specific mutations of Z-fragment of protein A in their interaction with a Fc fusion protein¹⁷. Two mutants (L17D and K35A) both had ten-times lower affinity for Fc binding compared with the wild-type protein domain. The L17D mutant had a ten-times lower

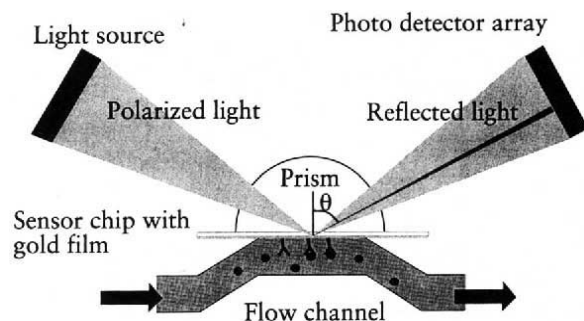


FIG. 1 The configuration of the surface plasmon resonance detector, sensor chip and integrated fluidic cartridge in the BIAcore system. The detector works with polarized light from a light-emitting diode, which is reflected in the gold film on the sensor chip and detected on a diode array. Surface plasmon resonance is observed as a decrease in light intensity for a specific angle of incidence. This angle changes with changes in the refractive index in the vicinity of the surface due to adsorption of large molecules on the immobilized ligand on the sensor chip.

association rate constant and K35A had a ten-times higher dissociation rate constant. These two examples show the importance of separating on and off rate constants. With a knowledge of the three-dimensional structure the information obtained can be used to relate structure to function.

Monoclonal antibodies directed against peptides from the coat protein of tobacco mosaic virus were analysed for binding both to different peptides and the protein¹⁸. The hydrophilic environment in the matrix is valuable here as conformational changes attributable to adsorption to plastic surfaces used in ELISA techniques are avoided. The kinetic analysis of mAb binding to a 25-amino acid peptide of the protein showed heterogeneous binding during association, in contrast with homogenous binding to a shorter peptide from the same sequence. The differences in affinity between peptides and protein could be ascribed to changes in association rate constants. Only kinetic analysis can separate these fine differences in the specificity of antibodies.

The effect of differences in peptide sequence in relation to the kinetic properties of mAb binding has been reported by Altschuh *et al.*¹⁶, as well as the binding of mAbs to whole virus particles^{19,20}.

The kinetic analysis of mAbs of IgG- and IgM-type against tetanus toxoid antigens illustrates the effect of multiple binding²¹, and a study of anti-hen lysozyme antibody, D1.3, and recombinant antibody fragments has been performed by the same research group²².

Epitope mapping. Epitope mapping of human immunodeficiency virus-1 (HIV-1) core protein p24 by pairwise mapping of 30 different mAbs using the BIAcore system has been performed by Fägerstam *et al.*¹¹. Both identification of peptides inhibiting the mAb binding and multiple binding of several mAbs in sequence forming a hexamolecular complex on the surface were shown (see Fig. 3). This illustrates the potential use of the system for the analysis of the formation of large functional complexes. The open structure of the surface matrix makes the formation of large complexes possible.

Recombinant antibodies. Kinetic analysis using BIAcore has been particularly useful in the rapid development of recombinant antibody technology. Affinity maturation of antibody fragments from a naive library by chain shuffling was found to be dominated by a decrease in dissociation rate constants¹³. Antibody fragments derived from a human naive library²³ characteristically had both high association and high dissociation rate constants²⁴. The system has also been used as a specific detector for analytical gel filtration showing the formation of active dimers of single chain Fv antibody fragments.

New application areas

The interaction of HIV-1 gp120 and a mAb with soluble CD4 has also been analysed with BIAcore¹². In application notes, research

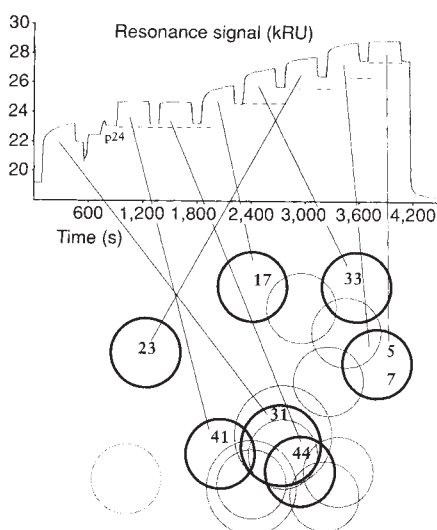


FIG. 3 Formation of a hexamolecular complex of five mAbs and HIV-1 core protein p24 on the sensor chip by adsorption of mAbs from tissue culture medium. The mAbs were selected from the epitope map illustrated in the lower part of the figure. The mAbs were sequentially injected over antigen bound to surface-immobilized mAb¹¹ (reprinted from ref. 11 by permission of John Wiley & Sons Ltd).

groups report examples of equilibrium and kinetic analysis of epidermal growth factor (EGF) and EGF receptor. The soluble receptor binding to immobilized EGF has high on and off rate constants. Other examples are the interaction of different SH2 domains with an immobilized phosphotyrosin peptide derived from platelet-derived growth factor receptor.

An application area under development is protein-DNA interactions with the immobilization of DNA through biotinylated nucleotides on streptavidin surfaces.

The technology, real-time biospecific interaction analysis, was originally developed using mAb-antigen interactions as model systems but is now rapidly being transferred to other areas of biological significance due to the general approach for label-free analysis of binding reactions. Although, to date, most scientific reports have been from academic research laboratories, the technology is now finding most applications in the pharmaceutical industry, both in research and in the quality control of recombinant proteins to be used as therapeutic agents. □

Magnus Malmqvist is at Pharmacia Biosensor AB, S-751 82 Uppsala, Sweden. For more information, fill in reader service number 100.

- Liedberg, B., Nylander, C. & Lundström, I. *Sensors Actuators* **4**, 299-304 (1983).
- Stenberg, E., Persson, B., Roos, H. & Urbaniczky, C. *J. Colloid Interface Sci.* **143**, 513-526 (1991).
- Löfås, S. & Johnsson, B. *J. chem. Soc., Chem. Commun. Issue* **21**, 1526-1528 (1990).
- Johnsson, B., Löfås, S. & Lindqvist, G. *Analyt. Biochem.* **198**, 268-277 (1991).

- Sjölander, S. & Urbaniczky, C. *Analyt. Chem.* **63**, 2338-2345 (1991).
- Jönsson, U. *et al. BioTechniques* **11**, 620-627 (1991).
- Jönsson, U. & Malmqvist, M. in *Advances in Biosensors* (ed. Turner, A.) 291-336 (JAI Press, San Diego, 1992).
- VanCott, T. C., Loomis, L. D., Redfield, R. R. & Birk, D. L. *J. Immunol. Meth.* **146**, 163-176 (1992).
- Löfås, S., Malmqvist, M., Rönnerberg, I. & Stenberg, E. *Sensors Actuators* **B5**, 79-84 (1991).
- Fägerstam, L. G., Frostell-Karlsson, Å., Karlsson, R., Persson, B. & Rönnerberg, I. *J. Chromatog.* **597**, 397-410 (1992).
- Fägerstam, L. G. *et al. J. molec. Recogn.* **3**, 208-214 (1990).
- Brigham-Burke, M., Edwards, J. R. & O'Shannessy, D. J. *Analyt. Biochem.* **205**, 125-131 (1992).
- Marks, J. D. *et al. Bio/Technology* **10**, 779-783 (1992).
- Karlsson, R., Michaelsson, A. & Mattsson, L. *J. Immunol. Meth.* **145**, 229-240 (1991).
- Karlsson, R., Altschuh, D. & Van Regenmortel, M. H. V. *Structure of Antigens* (ed. Van Regenmortel, M. H. V.), Vol. 1, Ch. 7, 127-148 (CRC Press, Boca Raton, Florida, 1992).
- Altschuh, D., Dubs, M. -C., Weiss, E., Zeder-Lutz, G. & Van Regenmortel, M. H. V. *Biochemistry* **31**, 6298-6304 (1992).
- Cederger, L. *et al. Symp. Structure and Function of Mutated Proteins*, Florence, Italy, 25-30 Aug., 1991, p.18.
- Zeder-Lutz, G. *et al. Molec. Immunol.* **30**, 145-155 (1993).
- Dubs, M. -C., Altschuh, D. & Van Regenmortel, M. H. V. *Immunol. Lett.* **31**, 59-64 (1991).
- Dubs, M. -C., Altschuh, D. & Van Regenmortel, M. H. V. *J. Chromatog.* **597**, 391-396 (1992).
- Malmborg, A. -C., Michaelsson, A., Ohlin, M., Jansson, B. & Borrebaeck, C. A. K. *Scand. J. Immunol.* **35**, 643-650 (1992).
- Borrebaeck, C. A. K. *et al. Bio/Technology* **10**, 697-698 (1992).
- Marks, J. D. *et al. J. molec. Biol.* **222**, 581-597 (1991).
- Griffiths, A. D. *et al. EMBO J.* (in the press).

ADVERTISEMENTS

PEPTIDE SYNTHESIS SUPPORTS

New polydimethylacrylamide & polystyrene supports in bulk quantities

- * Reproducible high quality
- * High loadings

Tel (413) 253 9554

Fax (413) 253 2476



Polymer Laboratories
SEPARATION SCIENCE DIVISION

Reader Service No.42

Highest Quality Oligonucleotides

Synthesised, purified and shipped in 3 days

£2.75 per base

(All inclusive)

0.2 µM scale (approx. 500 µg)

- Antisense (completely or partially sulphurised)
- Extensive range of labelling and modifications
- Full design and analysis facilities
- Gene cassettes, bulk order and contract discounts

Molecular Medicine Unit

King's College School of Medicine & Dentistry

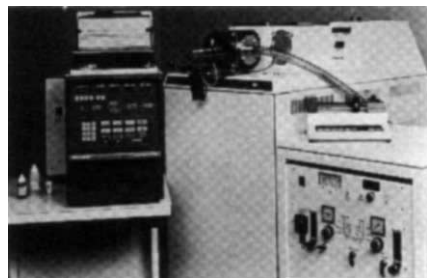
Tel. 071 - 326 3126 FAX 071 - 733 3877

Reader Service No.56

Protein and peptide product profile

Sequence analysis software, structure prediction programs, columns and media, and a range of products for glycobiology research — new products for a new year.

BECKMAN Instruments announces the release of the P/ACE 2100 System for analysis by **capillary electrophoresis combined with mass spectrometry (MS)** (*Reader Service No. 101*). When combined, Beckman says that these methods provide a useful analytical tool for the analysis of carbohydrates, nucleic acids, peptides, pharmaceuticals and proteins. Analyses using the P/ACE 2100



Beckman's CE system for use with MS.

system are designed to provide high-resolution separations rapidly with only a few nanolitres of sample. With MS as a detector for the P/ACE system, the company says that high-resolution separations are coupled with accurate molecular weight determinations, molecule identification and 10^{-11} sensitivity. The combination of CE and MS is particularly suited to the analysis of complex matrices as the resolving power of CE yields sharp peaks, which allows a greater accuracy in analysis by MS. In addition, analysis by MS/MS on a triple quadrupole mass spectrometer can provide sequence information for biopolymers such as oligosaccharides and peptides. Analytes eluting from the CE column are ionized via electrospray ionization, which is a 'soft ionization' technique capable of producing charged molecules that can be readily analysed with a single or triple quadrupole MS system. Beckman's P/ACE system can be interfaced to the Finnigan Mat SSQ 710 and TSQ 700 quadrupole mass spectrometers.

Promega's ProtoBlot II AP **western blotting systems** have now been improved to include the Western Express rapid blotting protocol (*Reader Service No. 102*). The Western Express protocol is designed to assist in the custom development of a blotting procedure to optimize detection time for the sensitivity of detection required. Promega says that proteins or other macromolecular antigens immobilized on nitrocellulose or PVDF membranes can be detected rapidly using the appropriate alkaline phosphatase-conjugated second antibody: dot blots can

be performed in 30–45 min and western blots in 1–2 h. The systems also incorporate a premixed, ready-to-use Western Blue Stabilized Substrate for alkaline phosphatase (NBT/BCIP), Tween 20 and Blot-Qualified BSA for optimizing blocking. ProtoBlot II AP Systems with the Western Express protocol contain sufficient reagents to process twenty (10×15 cm) membranes for the detection of human, mouse and rabbit antibodies.

PharMingen has expanded its line of **monoclonal antibodies specific for mouse cell surface antigens** with several new specificities, including C-kit, CD26 (DPP IV), Forssman antigen, interleukin-2 receptor β chain (p75), Mac-3 and Sca-1 (*Reader Service No. 103*). These antibodies are supplied both purified and conjugated and have been optimized for use in flow cytometry.

■ Sequence analysis software

Release 2.2 of GeneWorks, the **nucleic acid and protein sequence analysis software** package for the Apple Macintosh computer



GeneWorks 2.2 with PCR primer design.

available from IntelliGenetics/Betagen, includes new primer design functionality that is designed to take the guesswork out of polymerase chain reaction (PCR) experiments and new database searching capabilities (*Reader Service No. 104*). GeneWorks allows the user to visualize and manipulate sequence data to find the best possible PCR primer. It also produces reports that are highly visual and informative. GeneWorks will evaluate primers it designs, as well as primers suggested by the user. Parameters for primer design include: location, length, and annealing temperature of the amplified region; T_m and percentage GC; control over nucleotide runs and GC clamps; automatic exclusion of snapback, primer-dimer and other potentially dangerous interactions. GeneWorks 2.2 also includes a new translated nucleic acid database search. This search compares a

protein sequence against the GenBank or EMBL databases by translating the databases 'on the fly' in all six reading frames.

Hitachi Software Engineering has released version 2.0 of MacDNASIS, the company's **DNA and protein sequence analysis software** program for Apple Macintosh computers (*Reader Service No. 105*). The program provides a comprehensive set of DNA and protein sequence analyses, including sequence input and editing, primary and secondary structure prediction, 'shotgun' fragment assembly, restriction mapping, and similarity search/retrieval of genetic information from the GenBank, EMBL, NBRF-PIR and SWISS-PROT databases supplied on CD-ROM. Version 2.0 provides new structure prediction functions and improved dot-matrix homology, protein secondary structure graphics and RNA folding functions. The program also allows researchers to interpret visually results of various analyses by displaying data graphically. MacDNASIS costs \$1,950 (US).

■ Getting a grip

EM Separations has introduced a new integrated system for protein separations, which features interchangeable **cartridges prepacked with tentacle ion exchange media** and reusable Superformance glass columns (*Reader Service No. 106*). Tentacle ion exchangers are suitable for separating proteins, providing a two-to-four times greater capacity with better recovery, higher resolution and scale-up reproducibility, EM



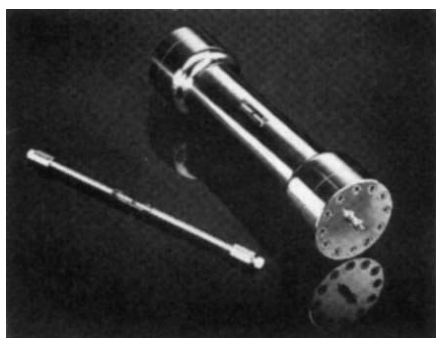
Tentacle ion exchange media and reusable glass columns.

Separations says. In addition, the design of the Superformance glass columns ensures zero dead volume, which results in sharper peaks. Cross contamination in multiple separations is also minimized.

With Fractogel EMD Chelate and Fractogel EMD TAAF, E. Merck now offers **tentacle carriers for affinity chromatography** (*Reader Service No. 107*). The affinity ligands are bound to the matrix via hydrophilic polymer chains, which can move freely in space and, therefore, adapt to the sterical structure of the bound protein. The company says that the spacer significantly reduces unspecific interactions with the carrier matrix. The ligands used are chemically stable and are designed to ensure highly reproducible separations, even after many purification and regeneration steps. The chelate type works with immobilized metal ions: the choice of ion is up to the user and should be determined according to the specificity of the biopolymer to be isolated. The thiophilic TA AF type is intended for the isolation and purification of immunoglobulins such as IgG and IgM. It can be used for analytical and preparative applications and can replace protein-A and protein-G affinity gels.

■ Columns and media

To complement its range of **polymer-based HPLC columns and media** for analytical and semipreparative scale separations, Polymer Laboratories has introduced large (50- and 100-mm ID) columns in 150- and 300-mm lengths, which are designed specifically for multigram biomolecule



Polymer-based HPLC columns and media.

purification (*Reader Service No. 108*). The new columns are machined from a solid blank of 316 stainless steel and are designed for operating at pressures of up to 3,000 p.s.i. which, for the high-resolution, reversed-phase purification of peptides, offers greater throughput. This new column hardware is packed with the company's polymeric media for reversed-phase (PLRP-S), ion exchange (PL-SAX/PL-WAX), and gel filtration (PL-GFC), for the purification of biomolecules, including peptides, proteins and DNA over a pH range of 1–14. The range of preparative- and process-scale media is available in 10-, 20- and 60- μ m particle sizes.

Pharmacia Biotech has introduced Superdex 200 HR 10/30, a **pre-packed column** for the high-resolution separation of peptides and proteins by gel filtration (*Reader Service No. 109*). Superdex 200 HR



Superdex 200 HR 10/30 prepacked columns for high-resolution gel filtration.

10/30 delivers optimal resolution in the molecular weight range of 10,000–600,000 daltons. It can be used for any application of gel filtration, such as monitoring changes in molecular size, determining molecular weight, or as a step in a purification scheme, when separating monoclonal antibodies, for example. Superdex 200 HR 10/30, like its lower molecular-weight range counterpart, Superdex 75 HR 10/30, is supplied prepacked for added convenience and reproducibility. Both columns run analytical and semipreparative applications in the microgram-to-milligram range on FPLC and HPLC systems.

PROSEP-G is a **protein-G affinity adsorbent** developed by Bioprocessing (*Reader Service No. 110*). The protein G is immobilized on porous glass, which, the company says, is more rigid and durable than conventional polymeric matrices. PROSEP-G will not only bind antibodies that bind to protein-A affinity matrices but also antibodies that may be difficult to bind to protein A, such as rat IgG_{2b} and human IgG₁. It can also be used for the purification of Fab fragments.

■ Sequence-structure analysis

PPSP (Protein and Peptide Structure Prediction) from European Scientific Software is a **program that can be used to model peptides** using a combination of sophisticated modules to predict peptide structures (*Reader Service No. 111*). The program is composed of several modules, which allow the structure of certain regions of the peptide to be identified. These domains are determined by a mixture of matching sequences with, for example, known structural-forming sequences, and by using a database containing the known optimized energy conformations of all tri-peptides, to generate a few most likely structures. This process will typically generate, at most, a few hundred possible structures, which can then, in minutes, lead to a likely final structure, the company says. PPSP will run

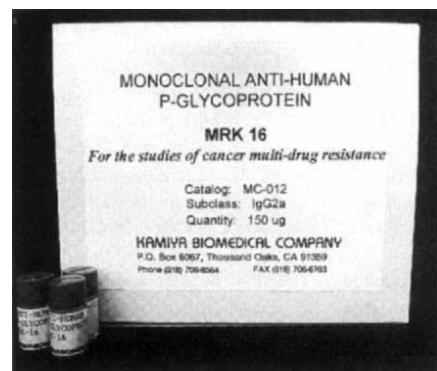
on most IBM PCs or compatible computers, although a 386 with a co-processor is recommended for the final modelling stages.

Biosym announces a collaboration to commercialize Profiles, a **program for analysing protein sequence-structure relationships** (*Reader Service No. 112*). The program tackles the problem of inverse protein folding, a significant tool in the development of structure-function relationships. Profiles was conceived and developed by David Eisenberg, James U. Bowie and Roland L  thy of the Molecular Biology Institute and the Department of Chemistry at the University of California, Los Angeles. It will be incorporated into Insight II, Biosym's three-dimensional (3-D) graphical interface, for release in the first six months of 1993. Biosym says that this integration provides the user with access to all of Biosym's programs, including the ability to move molecules between its other specialized rational drug design tools. The problem of what sequence folds into a given structure is addressed by Profiles. It can assign sequences to known 3-D structures, detect structural similarity among proteins with no detectable sequence similarity, be applied to mis-folded protein analysis, and can help determine the oligomeric state of proteins.

■ Glycobiology update

The new **chemical deglycosylation kit** from Oxford GlycoSystems is designed to help researchers remove bound carbohydrate from glycoproteins to facilitate characterization/sequencing or further study of the pure protein (*Reader Service No. 113*). The company has optimized reaction conditions to maximize the release of all carbohydrates, including O-linked oligosaccharides, while minimizing cleavage of peptide bonds. The new kit, GlycoFree, uses the optimized anhydrous trifluoromethane sulphonic acid reaction and includes all the necessary reagents to deglycosylate up to 1 mg of glycoprotein from up to 10 samples in less than 8 h.

Kamiya Biomedical has added a murine (IgG2a) **monoclonal antibody specific for a surface epitope of human P-glycoprotein**,



Anti-P-glycoprotein monoclonal antibody.

ADVERTISEMENTS

Severn Biotech Ltd.
CUSTOM SERVICES

- **OLIGONUCLEOTIDE SYNTHESIS**
High quality PROBES, PRIMERS, ANTISENSE OLIGOS!
Fast-Efficient Service offering loads of modifications:
BIOTIN, FLUORESCENCE, 5' & 3' labeling,
Phosphothioates Affinity Support/Oligos etc.
- **MULTIPLE PEPTIDE SYNTHESIS**
Epitope Mapping & Screening! We now supply 5mg
quantities of high purity custom peptides up to 15
residues in length, using Fmoc chemistry. With D or L
amino acids. These are supplied lyophilised in 4ml
vials, cleaved from solid phase & ready for use.

Tel: 0562 825286 Fax: 0562 825284
Unit 23 Huddbrook Enterprise Centre
Worcester Road, Kidderminster DY10 1HY

Reader Service No.26

**epitope
custom
peptides**

18 Windsor Terrace, Newcastle upon Tyne, UK, NE2 4HE.

HIGH QUALITY CUSTOM MADE PEPTIDES
COMPUTERISED EPITOPE PREDICTION
ANTIBODY PRODUCTION
PROTEIN SEQUENCING

PHONE OR FAX FOR A FREE QUOTATION
TEL: +44 (0) 91 222 7186
FAX: +44 (0) 91 221 0074

In association with Newcastle University Ventures

Reader Service No.24

PACEMAKERSM

The multiple peptide synthesis and conjugation service

- High purity (up to 95%).
- Quantities for a range of uses.
- Single or multiple peptides.
- HPLC trace with each peptide.
- Terminal modification at no extra charge.
- Conjugation to carrier proteins.

from **£80 per peptide**

AFFINITY Research Products Ltd.
GPT Business Park, Technology Drive,
Nottingham, NG9 2ND, UK.
Tel: +44 (0) 802 436100; Fax: +44 (0) 802 436300

Reader Service No.23

Clarification

The two fluorescent nucleic acid stains, TOTO and YOYO (or TOTO-1 and YOYO-1) mentioned in the product review "Stable dye-DNA intercalation complexes as reagents for high-sensitivity fluorescence detection" (*Nature* **359**, 859-860; 29 October 1992) are only two of eight dyes of this type invented by scientists at Molecular Probes, Inc. This series of dyes is available only from Molecular Probes at the following address: 4849 Pitchford Avenue, Eugene, Oregon 97402-9144, USA.

MRK 16, to its product line (*Reader Service No. 114*). Research has linked multidrug resistance to the overexpression of P-glycoprotein, a 170-kD membrane-bound protein. P-glycoprotein appears to act as an ATP-dependent efflux pump to remove drugs from the cell. Kamiya Biomedical feels that the MRK 16 monoclonal antibody will be of interest to researchers studying multidrug resistance in cancer. MRK 16 is prepared by immunizing a mouse with an adriamycin-resistant K562/ADM cell line that has been established from a human myelogenous leukaemia cell line, K562.

Ciba Corning has upgraded Glycomat, the company's fully automated, **low-pressure liquid chromatography system** (LPLC), to analyse haemoglobin variants in addition to glycosylated haemoglobin



Glycosylated haemoglobin and more on the 765 Glycomat from Ciba Corning.

fractions (*Reader Service No. 115*). Three different colour-coded LPLC microcolumns and associated reagent kits allow for a wide range of haemoglobin assays. Yellow microcolumns and reagents provide a determination of HbA1C in 10 min, even when haemoglobin variants are identified, Ciba Corning says. Blue microcolumns enable HbA1C assays in 5 min. When a variant is detected, a second analysis is triggered to quantify correctly HbA1C excluding any variant fractions. Green microcolumns and reagents offer an estimation of HbA2, a fast HbS screen and full Hb variant analysis. The fast HbS test provides a rapid method of screening for the presence of HbS, HbD and HbC.

The CarboPac MA1 **anion exchange column** from Dionex has been specifically designed for the analysis of weakly ionizable carbohydrates such as sugar alcohols (*Reader Service No. 116*). The column can be used to separate reduced mono- and disaccharides commonly found in food products as well as those found in physiological and glycoconjugate samples. Specifically designed to retain and separate reduced carbohydrates, the CarboPac MA1 separates sugar alcohols as their anions by high-performance anion exchange chromatography. Detection of these complex compounds has been optimized with the use of pulsed amperometric detection. This set-up offers several advantages: an improved selectivity when compared to metal-loaded resins, a higher

sensitivity than refractive index detection, sample derivatization is not required and separations can be carried out at ambient temperature.

Catalogues and custom services

Genzyme has introduced an improved Lectin Link kit, which enables the **detection and characterization of N-linked oligosaccharides on glycoproteins** (*Reader Service No. 117*). Details of the kit and the company's full range of glycoprotein processing enzymes and inhibitors can be found in the 1992 GlycAnalysis catalogue. Genzyme has also developed a liquid-stable reagent for the direct estimation of serum amylase. Based on a simplified, short oligosaccharide with the direct release of an improved chromogen, the liquid-stable reagent contains 2-chlor-4-nitrophenyl maltotriose (CNP3). Genzyme says that, unlike other amylase substrates, the reagent does not require additional exoglycosidases for colour formation. The new technology provides an amylase test format that gives good correlation with conventional methodologies, with no lag phase in the reaction and negligible interference.

A new, informative **wallchart** is available from Cambridge Research Biochemicals depicting aspects of solid-phase peptide chemistry (*Reader Service No. 118*). It provides a range of useful information, including molecular weights of amino acids, and commonly used protecting groups, linkers and coupling reagents. Preferred methodologies for both Boc and Fmoc chemistries are highlighted.

Epitope Custom Peptides is offering a **custom peptide synthesis service** to researchers in industrial, academic and medical research fields (*Reader Service No. 119*). Customers are offered a standard purity of greater than 80 per cent (95 per cent on request), a standard quantity of 20 mg, reversed-phase HPLC, amino acid analysis, and time of flight mass spectrometry. For an additional charge, Epitope will also conjugate peptide to carrier proteins. The company also offers a computerized epitope prediction service to enable researchers to establish protein configuration and epitopic sites. The charge for this service depends on the complexity of the protein and the computer time that is required. For researchers who are new to this field and who are interested in setting up and running a monoclonal laboratory, Epitope offers a handbook entitled "*The Epitope Guide to Successful Monoclonal Antibody Production: From Peptide to Product*". □

These notes are compiled by Diane Gershon from information provided by the manufacturers. For more information, fill in the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.